

THIS WEEK

EDITORIALS

LIGHTSPEED The continuing revolution in space communications **p.254**

WORLD VIEW Don't let commercial firms dictate scientific terms **p.255**

WALKIES! Ancient American dogs took land route not sea **p.256**



The benefits of brain mapping

Two huge projects have the potential to revolutionize neuroscience, as long as they don't drain money from other work and are monitored to keep them on target.

There are roughly as many neurons in the human brain as there are pages on the Internet, give or take a million or so. But brain wiring is orders of magnitude more complex than the web. "Think about it like this," says Konrad Kording, a neuroscientist at Northwestern University in Chicago, Illinois, who is interviewed in a News Feature on page 272 that describes the task ahead for two billion-dollar research programmes to understand the brain. "Whereas Internet pages only link to a couple of others in a linear way, each neuron links to thousands of others — and does so in a nonlinear way."

Neuroscientists know frighteningly little about the brain's complexity. They have sketched out the broad anatomy of the brain, and realize that individual functions — from deciding to lift a cup to feeling envy — are mediated by circuitry that crosses anatomical borders. They can examine the detailed electrical activity of small numbers of neurons. They can wield imaging technologies that show which brain areas are activated during defined tasks, such as viewing pleasant or unpleasant pictures. But those tiny (in brain terms) pieces of information have not led neuroscientists to the big picture: what we mean by human consciousness, what makes us our individual selves or why some people develop psychiatric disorders. Neuroscientists need to be able to join the dots — and there are a lot of dots.

Many scientists now believe that real progress on learning how the brain works can be made only through highly funded, interdisciplinary big science of the kind promised by US President Barack Obama's BRAIN Initiative (Brain Research through Advancing Innovative Neurotechnologies) and the European Commission's Human Brain Project, both launched this year. The first steps will be to develop technologies to map the brain in unprecedented detail, in terms of activity and anatomy — and to develop theoretical neuroscience to make sense of it.

The approach is intimidating. Each of the big programmes is expected to absorb US\$1 billion or more in the next decade (although not all of the money is yet in the bank). Some neuroscientists worry that pumping so much money into top-down programmes will reduce support for small, hypothesis-driven projects in individual labs. They are also concerned that the money could be wasted by pouring it into schemes that turn out to be unhelpful.

They are right to be nervous, and their concerns must be taken on board. The money pumped into the launch phases of the two big programmes is genuinely new — and the rest must be, too. As the programmes develop, they should not encroach on funding for research driven by individual investigators. And it is essential that the broad scientific community is involved in design and oversight of the big programmes, to ensure that the processes remain transparent, on track and based in the real world.

In some ways, the current tensions in the neuroscience community bear comparison to those surrounding the launch of the Human Genome Project in the early 1990s. New technologies had revolutionized molecular biology in the previous decade, but the outpouring of

data from the project could not address the big picture of how genes fit together to keep us healthy or make us ill. The data needed to be referenced to detailed genome sequences. Many in the community bewailed the centralization and industrial scale of the work. Scientists prophesied the end of individual-investigator-driven research. But in the end, the results served to promote individual projects — and have revolutionized our understanding of many diseases. No molecular biologist now regrets the centralized investment in the Human Genome Project, even

"Neuroscientists know frighteningly little about the brain's complexity."

if — perhaps even because — it unexpectedly showed that we are only partly controlled by the sequences of our DNA. Researchers are now pinning down the mechanisms through which gene expression is altered by our changing environment.

Similarly large gains are likely to emerge from the big brain projects, although they are not inevitable. The risk of failure is high, but the rewards of success will be great. In recent years, much of the pharmaceutical industry has pulled out of work on brain disorders, following many failures of candidate drugs in trials. There is an urgent need for new approaches to treating disorders that manifest in the young, such as schizophrenia or autism spectrum disorder, and those that begin in older age, such as Alzheimer's or Parkinson's disease. Reliable and detailed reference maps of the brain are likely to ground the research in reality, attracting industry back.

There is also a high-stakes cultural issue. As neuroscientists get a stronger grip on how the brain works, they will encroach ever deeper into the territory of philosophers. What does it mean to be human? The brain will contain the answer — and it won't be about surfing the Internet. ■

Active protection

Parents should vaccinate their children against human papillomavirus.

Scientists at the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, announced good news last month. The prevalence of key strains of disease-causing human papillomavirus (HPV) fell by 56% in US girls aged 14–19 years in the years after 2006, when a vaccine was added to the routine US immunization schedule for girls (L. E. Markowitz *et al.* *J. Infect. Dis.* **208**, 385–393; 2013).

This is a clear-cut vaccine success story. The decline represents a drop from more than 1 in 10 girls in this age range carrying the



Don't market stem-cell products ahead of proof

The controversy over an unproven stem-cell therapy in Italy highlights the dangers of doing translational medicine in reverse, argues Paolo Bianco.

Translational medicine is said to reflect a need to harness the huge wealth of scientific knowledge in biomedicine. In fact, it is a direct consequence of the globalized outsourcing of research and development by the pharmaceutical industry, resting on the creation of commercial enterprises within academia. A commercial drive in academia can, however, significantly alter scientific concepts in biology and medicine.

Mesenchymal stem cells (MSCs) provide a prime example of this. Decades of research on these cells, found in the bone marrow, show that they go on to form skeletal tissues such as bone, fat and cartilage, which they can also help to regrow and repair in the clinic. Yet companies have already emerged that market MSCs for a much broader range of applications. Against mainstream scientific evidence, these firms argue that the cells are veritable injectable drug stores.

This commercial creep has reached the pages of authoritative scientific journals, with articles suggesting that intravenously infused MSCs can be used as a single agent to mute or cure a long list of unrelated diseases in multiple organs, regardless of their cause and nature. Notably, these include terminal neurodegenerative diseases, strokes and heart attacks. These are extraordinary claims that would require extraordinary evidence, which, in my view, does not yet exist. The very concept of MSCs has become divorced from that of a stem cell found in bone marrow. The scientific literature now contains two conflicting descriptions of these cells — one based on science, another on commerce.

Industry has not yet generated conclusively proven medicinal products or major novel technologies to better harness the biology of MSCs. However, commercial interest has profoundly influenced the definition of these cells (and of their clinical potential) within the scientific community. This is translational medicine in reverse. Commercial products have been converted into scientific concepts. It highlights an important dark side of the commercialization of science.

The marketing of MSCs as a cure-all is no coincidence — they have long been credited with potential performances that are beyond their biological limits. A decade or so ago, they were promoted as an ethical alternative to pluripotent stem cells derived from human embryos. They lost that unique selling point with the arrival of a technique to genetically reprogram adult cells into pluripotent cells. Suddenly, MSCs became 'pluri-effective' through intravenous infusions and release of chemical factors. Yet intravenously infused MSCs die rapidly and are quickly cleared from the body. As shown by 50 years of *in vivo* experiments, locally transplanted MSCs form bone. They do so, data show, even if transplanted into the heart or brain.

How MSCs would mute or cure unrelated

diseases has never been fully explained, in my view. Proponents say that the cells would restore brain function by nurturing cells with chemical factors. This remains unproven. With no reliable preclinical rationale, trials of MSCs can never be anything other than inconclusive. The only winners are the firms wishing to sell the therapies, which add the trial details — if not the results — to their marketing brochures.

Some 300 clinical trials on MSC infusions have been, or are being, conducted worldwide. Their mere initiation, paradoxically, is used to suggest that intravenously infused MSCs can cure multiple unrelated diseases, which (to my knowledge) is not proven at this time. These statements, and the trials that fuel them, represent a new kind of advertisement within science. They can distort science and medicine, mislead the public, create illusions for patients, sabotage health-care systems and, above all, obstruct rather than accelerate the growth of science and the

development of medicine from it.

This is a worldwide problem, highlighted by current events in Italy. A Brescia-based organization called the Stamina Foundation is promoting an unproven MSC therapy to vulnerable patients, including children with lethal neurological diseases. This has in effect forced the Italian government to test the therapy in a government-funded clinical trial, to cope with a media-generated social crisis and the risk of infringing European regulations on stem-cell therapies. Last week, *Nature* called for the trial to be scrapped (see *Nature* **499**, 125; 2013) as evi-

dence emerged of terminal flaws in the biology behind the purported cure. Stamina is backed by companies and a lobbying organization called the Cure Alliance, which has offices in Milan and Rome.

Central to the agenda of those who promote unproven therapies is an attack on the regulations surrounding such treatments, as well as the regulatory bodies that enforce them. Bone-marrow transplantation, some say, would never have been developed under today's stringent regulations. But bone-marrow transplantation was never a commercial product, and it developed when no one was out to sell stem cells directly to patients and ahead of proof.

Translating science into effective medicine cannot be based on indiscriminate development of commercial products. The push to fund commercial science and the seep of commercial descriptions of natural objects within academia can severely affect science, medicine and the economy. Claiming the right to market products ahead of proof of efficacy can only bring ineffective products to market, degrade medicine and impoverish all except, perhaps, the fortunate sellers. ■

Paolo Bianco is a stem-cell biologist, and professor and director of anatomic pathology at Sapienza University of Rome, Italy. He is an editor of Stem Cell Research. e-mail: paolo.bianco@uniroma1.it

**THE ONLY
WINNERS
ARE THE FIRMS
WISHING TO SELL THE
THERAPIES.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/pngxpt

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

MEDICAL MICROBIOLOGY

HIV disrupts gut bacteria

The guts of people with HIV are enriched with microbes associated with inflammation — even if patients are on antiviral therapy.

Mike McCune and Susan Lynch of the University of California, San Francisco, and their colleagues characterized microbial communities associated with chronic inflammation in patients with HIV. Communities with a higher abundance of Proteobacteria (including *Escherichia*, *Pseudomonas* and *Salmonella* species) and a lower abundance of Bacteroidetes were linked with higher levels of inflammation and with increased activity in a metabolic pathway whose products regulate T cells. These results indicate that microbes in the gut mucosa may influence progression of HIV disease.

Sci. Transl. Med. 5, 193ra91 (2013)

GENETICS

Old origins for New World dogs

Ancestors of American dog breeds may have walked across the Bering Strait rather than being brought across the ocean.

Peter Savolainen at the KTH Royal Institute of Technology in Solna, Sweden, and his colleagues collected mitochondrial DNA from blood or cheek cells of 347 individuals of indigenous American breeds, such as Chihuahuas,

Peruvian hairless dogs (pictured) and Canadian Eskimo dogs. The authors compared this DNA to that of modern European and East Asian dogs, as well as to 24 preserved New World specimens dated to long before Christopher Columbus set sail.

Modern native American canines, such as the feral Carolina dog, still resemble their pre-Columbian American counterparts, having no more than



30% European heritage. These modern breeds descended from Asian breeds brought by human migrants as long as 15,000 years ago, the authors suggest.

Proc. R. Soc. B 280, 20131142 (2013)

MATERIALS SCIENCE

Liquid metal printed in 3D

Wires, fibres and elaborate stacks of droplets can be printed in liquid metal.

Most three-dimensional (3D) printing uses molten plastics that cool and harden.

This cooling period changes the plastics' mechanical properties, which limits the shapes that can be created. A team led by Michael Dickey at North Carolina State University in Raleigh produced patterns in liquid metal by extruding a gallium–indium alloy through a 3D printer's nozzle at room temperature. On exposure to air, the material instantly formed a roughly nanometre-thick oxide skin, which held the liquid in shape. This layer was sticky and so allowed the team to stack droplets into complex constructions (pictured).

Although the structures were quite weak, the wires



CONSERVATION SCIENCE

Conserved coasts curb storm damage

Conserving the reefs and vegetation that buffer the US coast from waves might reduce by half the number of residents most at risk from storm surges and sea-level rise.

To work out where natural habitat provides the best defence, a team led by Katie Arkema at Stanford University in California used projections of future sea levels to estimate how vulnerable people and property would be to coastal hazards with and without intact natural habitats. Hazard indices calculated for every

square kilometre of the US coastline showed that ecosystems had the greatest protective impact in Florida, New York and California.

In places where natural habitats most reduce risks, conservation or restoration should be considered alongside expensive engineering projects for coastal defence, the authors suggest.

Nature Clim. Change <http://dx.doi.org/10.1038/nclimate1944> (2013)

For a longer story on this research, see go.nature.com/i8owbq

SEVEN DAYS

The news in brief

POLICY

Climate partners

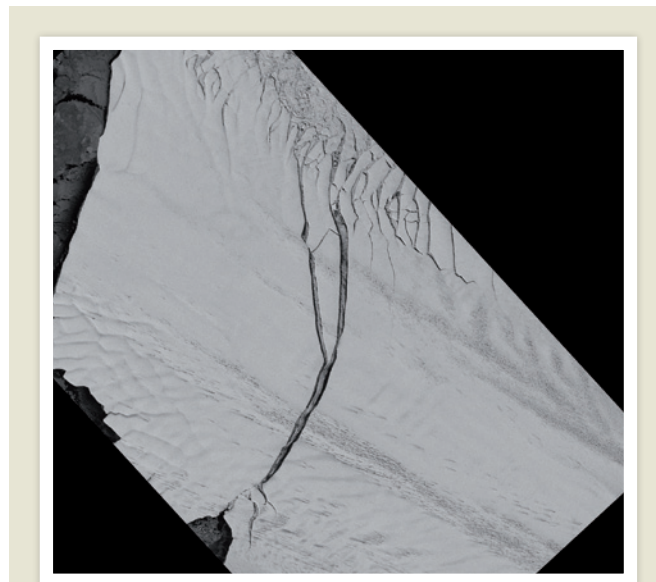
The United States and China will cooperate on reducing emissions from heavy-duty vehicles as part of a broad set of climate initiatives they announced on 10 July. Building on an earlier US–Chinese agreement to curb emissions of hydrofluorocarbons — powerful greenhouse gases used as refrigerants — the partnership aims to boost carbon capture and storage, increase building efficiency, improve greenhouse-gas monitoring and build more efficient electric grids. The two countries will develop plans by October 2013 to implement the initiatives.

Biofuels brake

Biofuels made from food crops are on course to be curbed in Europe after an 11 July vote by the environment committee in the European Parliament. Politicians moved to limit food-based fuels to 5.5% of transport fuel by 2020, after scientists warned that production of some biofuels drives land clearance that can lead to greater greenhouse-gas emissions than from fossil fuel (see *Nature* **499**, 13–14; 2013). Fuel suppliers would also have to report emissions due to changes in land use. The European Parliament votes on the issue in September.

Italian animal work

Proposed legislation in Italy to tighten restrictions on animal research has moved forward, entering the Chamber of Deputies on 11 July. The proposal, approved on 4 July by the Senate, would ban research involving addictive drugs or transplants of living cells from other species. It would also stop the breeding of dogs, cats and primates in Italy for research, and require



Giant iceberg cut adrift

Antarctica's Pine Island Glacier shed a massive iceberg last week (to the left of the crack), breaking along a deep rift that researchers had been monitoring since 2011. The 'calving' event dumped a 720-square-kilometre chunk of ice into the Amundsen Sea off western Antarctica, the German Aerospace Center (DLR) reported on 9 July, on the basis of images collected by its TerraSAR-X satellite. Pine Island, one of Antarctica's fastest-moving ice streams, previously calved large slabs of ice in 2001 and 2007. Researchers hope that careful charting of the latest event will improve their understanding of how climate change may affect ice loss by calving.

anaesthetic for any procedure causing mild pain in animals (such as giving injections). Concerned scientists say that the restrictions could halt important biomedical research, and fear that Italy is becoming increasingly hostile to animal studies.

Indian power

India's plans to expand its nuclear power capacity are moving ahead, despite public opposition following the Fukushima nuclear accident (see *Nature* <http://doi.org/ckcr86>; 2011). The first of two Russian-built reactors at Kudankulam in Tamil Nadu achieved a sustained chain

reaction on 13 July — a first step towards producing power. Work had stalled after protests in 2011, but in May, India's Supreme Court dismissed the concerns. The second reactor is expected to start up within eight months; the reactors will have a combined electricity generation capacity of 2,000 megawatts.

FUNDING

NIH funding

The US National Institutes of Health (NIH) received a tentative boost on 11 July when the US Senate committee on government spending approved a US\$31-billion

budget plan for the agency — nearly \$2 billion more than the NIH received this year. Although the plan must still be voted on by the full Senate, NIH supporters cheered the proposed increases. The plan includes \$84 million in new funds for Alzheimer's disease research at the NIH's National Institute on Aging. See go.nature.com/qkxiar for more.

BUSINESS

Myriad back in court

One month after the US Supreme Court invalidated gene patents held by Myriad Genetics of Salt Lake City, Utah, the company has sued two competitors for infringing different patents on tests for the cancer-related genes *BRCA1* and *BRCA2*. On 9 July, Myriad brought a lawsuit against Ambry Genetics of Aliso Viejo, California, then filed another the following day against Gene by Gene in Houston, Texas. Both firms had announced that they would provide *BRCA* testing in the wake of the Supreme Court ruling (see *Nature* **498**, 281–282; 2013).

Alzheimer's retest

Eli Lilly, a pharmaceutical company based in Indianapolis, Indiana, announced on 12 July that it will focus on patients with mild Alzheimer's disease in a forthcoming clinical trial of the drug solanezumab. The antibody-based drug, which targets the amyloid- β protein, will be tested in a large-scale phase III trial. In 2012, the company reported lacklustre results in two previous trials, which pooled data from patients with mild and moderate forms of Alzheimer's. But secondary analysis hinted that the drug might help patients with mild forms of the disease.

DLR

RESEARCH

Red rover

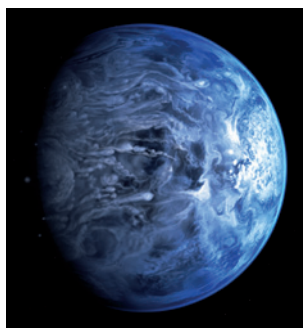
NASA's next Mars explorer will be a leaner, meaner version of the Curiosity rover, with one major upgrade: the ability to store rock and soil samples for return to Earth, the agency said on 9 July. The vehicle, planned for launch in 2020, will cost about US\$1.5 billion. NASA proposed the mission in December, less than a year after it disappointed planetary scientists by pulling out of Europe-led Mars missions planned for 2016 and 2018. See go.nature.com/au5ewk for more.

Research restart

Research on the rinderpest virus is set to resume after being off limits since 2011, when the deadly cattle disease was eradicated. The ban was enacted as a temporary measure to safeguard against accidental or intentional release of the virus. The moratorium was lifted on 10 July and replaced by an international oversight system. See page 264 for more.

True blue planet

Using the Hubble Space Telescope, astronomers have discovered the deep blue hue of exoplanet HD 189733 b (pictured in an artist's impression) — the first planet beyond the Solar System



to have its colour directly measured. Discovered in 2005, the planet orbits a star about 19 parsecs away in the Vulpecula, or Fox, constellation. At Hubble's optical resolution, light from the planet and its star typically blend together. But researchers found that the amount of blue light decreased when HD 189733 b ducked behind its star. See go.nature.com/pyze44 for more.

EVENTS

Negligence claim

On 11 July, relatives of two biologists murdered in 2010 at the University of Alabama, Huntsville, asked a judge to add the former university president as a defendant in their ongoing lawsuit to collect civil damages for the killings. Lawyers for the victims' families say that three months before the murders, David Williams, along with then-provost Vistasp Karbhari,

sought police protection for themselves from biologist Amy Bishop, without warning or protecting others. On 12 February 2010, Bishop shot six colleagues during a faculty meeting, killing three.

Economy rocket

The European Space Agency (ESA) announced the design of its next rocket, Ariane 6, on 9 July. The rocket will have less lifting capacity than the current model: taking only one satellite per launch instead of two. ESA's choice of a more cost-effective design was influenced by competition from rockets abroad, notably the Russian Proton launcher. Ariane 6 is expected to go into service during the next decade, at a cost of €70 million (US\$91 million) per launch.

PEOPLE

Theft plea

A former researcher at the Medical College of Wisconsin, Milwaukee, pleaded guilty last week to accessing a protected computer without authorization and taking information worth in excess of US\$5,000. Hua Jun Zhao, from China, had been under investigation for allegedly stealing patented cancer-research material, with plans to provide the material to Zhejiang University in Hangzhou, China. Zhao is

COMING UP

19 JULY

NASA's Cassini spacecraft turns to image Saturn and its entire ring system, while also capturing a picture of Earth from 1.44 billion kilometres away.

20–24 JULY

In Kagoshima, Japan, researchers discuss volcano monitoring and forecasting at the Scientific Assembly of the International Association of Volcanology and Chemistry of the Earth's Interior.

go.nature.com/9cbgun

21–25 JULY

Current research in viral ecology, biofilms and zoonotic pathogens is on the agenda at the 5th Congress of European Microbiologists in Leipzig, Germany.

go.nature.com/h7nzes

scheduled to be sentenced on 6 August, and could face a \$250,000 fine and up to five years in prison.

California head

The United States' flagship public university system may soon have its first woman president. On 12 July, the University of California announced its nomination of Janet Napolitano, US secretary of homeland security and former governor of Arizona, to run the ten-campus system. If approved, Napolitano, who holds a law degree, would break into a position long dominated by academics. The university's board of regents will vote on her nomination on 18 July. Napolitano would replace Mark Yudof, who has served since 2008.

► NATURE.COM

For daily news updates see:

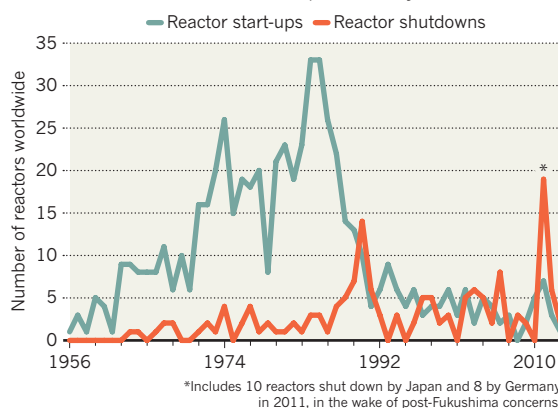
www.nature.com/news

TREND WATCH

Nuclear power plants generated 2,346 terrawatt hours of electricity in 2012, a decline of nearly 7% from 2011 and nearly 12% from the all-time peak in 2006, according to an analysis sponsored by the Berlin-based Heinrich Böll Foundation and the Greens–EFA political party in Europe. Much of the decline was due to power plant closures in Japan following the Fukushima nuclear accident in 2011, but the industry is also struggling to maintain and replace older plants.

GLOBAL NUCLEAR SLOWDOWN

Despite talk of a second nuclear renaissance, nuclear power plant shutdowns have outnumbered start-ups in recent years.



NEWS IN FOCUS

BIOTECHNOLOGY Sleuthing in a bid to solve a GM wheat whodunnit **p.262**

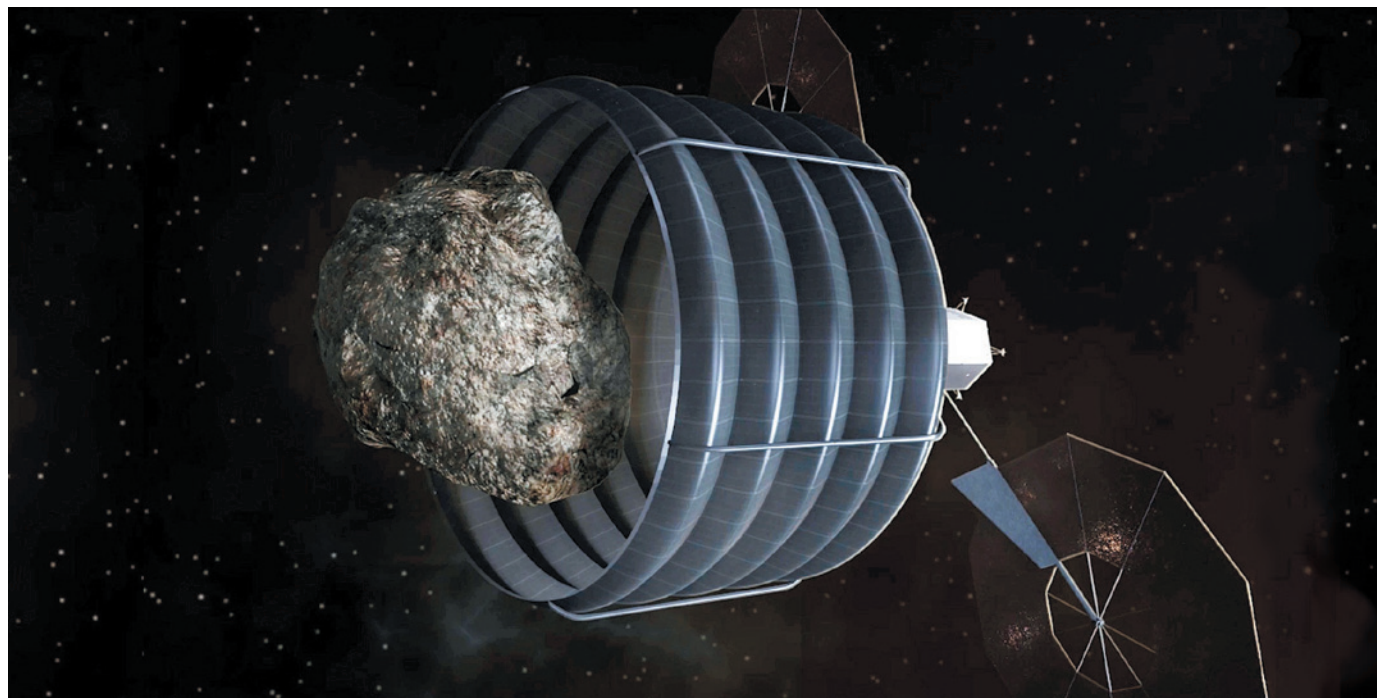
PHARMACEUTICALS New uses for old drugs in translational gamble by the NIH **p.263**

SPACE Lasers set to open up communications floodgates **p.266**

EDUCATION Lab practicals and politics, taught online **p.268**



NASA



NASA's asteroid-capture spacecraft (concept illustration) would envelope a 10-metre-wide rock in an inflatable bag and tug it close to the Moon.

SPACE

Asteroid plan looks rocky

NASA mission to retrieve a small space rock could be tripped up by lack of candidates.

BY ALEXANDRA WITZE

NASA has plans to grab an asteroid and tow it near to the Moon for astronauts to visit. But finding a target space rock may be the agency's biggest challenge. Only a handful of known asteroids would suit NASA's requirements, and current sky surveys are not tuned to find many more candidates.

"There's great scepticism, among both the science community and the public, that this can actually be pulled off," says Jim Bell, a planetary scientist at Arizona State University in Tempe, who attended a workshop on 9 July at the US National Academy of Sciences in Washington DC. The workshop marked the first chance for asteroid scientists to voice their doubts to NASA officials.

US President Barack Obama proposed the

asteroid initiative in April, as part of his 2014 budget request to Congress. If the mission gets funded, NASA would expand its surveys for large, hazardous space rocks so that the agency might find an asteroid that is small enough, solid enough and on a trajectory that would allow it to be met by a spacecraft launching as early as 2017.

Details of the mission remain fuzzy, but at the workshop, scientists revealed the first hard numbers on how many asteroids might be suitable to snatch. Of the more than 10,000 known near-Earth asteroids, 370 are small enough for capture, at roughly 10 metres across, says Paul Chodas, an asteroid tracker at NASA's Jet Propulsion Laboratory in Pasadena, California. But of those 370, only 14 have a suitable orbit. Just four have been studied well enough for scientists to know something about the

bodies' surface texture and spin rate. NASA wants to find a cohesive body spinning at less than two revolutions per minute, to minimize the risk of any damage to the spacecraft.

If NASA picks up the pace of its sky surveys, astronomers should find at least 15 more 10-metre targets over the next three to four years, says Chodas, and at least half of those are likely to be in the right orbit. But that would require the agency to boost funding to the survey programmes, which currently focus on finding rocks that are 140 metres across or larger. To spot 10-metre-class rocks, "we have to go fainter", says Timothy Spahr, director of the Minor Planet Center in Cambridge, Massachusetts, which catalogues near-Earth asteroid discoveries. Going fainter means, for instance, adding two more cameras to the Catalina Sky Survey, based at the University of Arizona ►

► in Tucson. The survey has discovered most of the small near-Earth asteroids so far. These cameras — planned for mid-2014 — would double the survey's field of view and increase the chances of catching a faint small object, says Spahr.

NASA also plans to restart the Wide-field Infrared Survey Explorer, a spacecraft that has been in hibernation since 2011, to hunt for more asteroids. Two facilities slated for Hawaii — the Pan-STARRS-2 telescope and the ATLAS telescope array — could also add to the effort.

The extra eyes are crucial because near-Earth asteroids often vanish from view within days of a sighting. On discovery, astronomers must quickly marshal follow-up observations with radar and infrared telescopes to learn more about the rock's size, shape and composition.

Engineers at NASA's Johnson Space Center in Houston, Texas, are already working on several mission designs for the robotic capture craft, which would need to grapple with an asteroid weighing around 400 tonnes — about the mass of the International Space Station (ISS). The xenon-ion thrusters used to get to the asteroid, and the inflatable rip-resistant bag used to capture it, are straightforward components. The bigger issue is finding a target and building the spacecraft in time for launch, says Andrew Thomas, an astronaut at Johnson who works on the project. "It's pretty brazen, if you think about it," he says.

A 2017 launch would allow the spacecraft to drag the target asteroid to a position near the Moon by 2021. In theory, astronauts could then visit the asteroid, peel back or cut the bag and take rock samples back to Earth.

NASA expects the mission to cost between US\$1 billion and \$2.6 billion. Agency officials are careful to emphasize that the goal is not science, but rather to give astronauts practice in going into deep space, where they must contend with harsher radiation environments than on missions to the ISS. "If our goal is to get to Mars in the 2030 time frame, we're going to have to be aggressive with this upfront stuff," says William Gerstenmaier, NASA's associate administrator for human exploration and operations.

It is not yet clear whether Obama's plan will be paid for by Congress, which has been locked in a battle with the administration over agency goals ever since NASA gave up former President George W. Bush's vision of returning to the Moon (see *Nature* **492**, 161–162; 2012). A bill proposed by a congressional subcommittee on space in the Republican-dominated House of Representatives would cancel all funding for the asteroid initiative. ■



Wheat in a test field at Oregon State University where scientists are helping to identify a rogue GM crop.

BIOTECHNOLOGY

Hunt for mystery GM wheat heats up

Investigators hope to track origins of the transgenic crop.

BY HEIDI LEDFORD

It has been nearly three months since an Oregon farmer discovered unapproved transgenic wheat in a commercial wheat field, triggering bans on imports of US wheat into Japan and South Korea. The harvest season has now begun, and with the contamination proving to be an isolated event, imports into South Korea have resumed.

But as an army of combines marches across the wheat fields of eastern Oregon, the mystery of the transgenic intruders is fresh in the minds of investigators at the US Department of Agriculture (USDA), who are trying to trace the plants' provenance to a particular research plot. Those close to the investigation say that the identity of the variety in question could emerge in the coming weeks, providing a much-needed breakthrough in this agricultural whodunnit. It could point to the cause of the release, which some have suggested could be activist sabotage. "We may never know who actually released it," says James Moyer, director of the Agricultural Research Center at Washington State University in Pullman. "But if they know the genotype of those plants, they will be able to narrow it down quite a bit."

Within a month of the discovery in May, USDA scientists had traced the origin of the plants to a line of herbicide-resistant 'Roundup

Ready' wheat called MON71800, developed by the agricultural company Monsanto, based in St Louis, Missouri. Monsanto killed the project in 2005 over farmers' worries that overseas customers would not buy US wheat if it contained transgenic varieties. No GM wheat has yet been approved to be grown commercially in the United States. The company says that all seed from the field trials — conducted on more than 400 hectares in 16 states (see 'Sifting for GM wheat') — was accounted for and either secured or destroyed.

Monsanto had shipped MON71800 seed to breeders around the country for crossing with commercial varieties optimized for each region's climate, day length and disease profile. Now, the USDA investigators are sifting through hundreds of markers to try to match the genetic signature of the contaminant Oregon wheat with one of the varieties from the 256 field tests registered with the USDA. An origin near Oregon could hint at an accidental escape; seeds from farther afield could mean that someone had intentionally saved seed and released it.

But pinning down the variety is difficult, says Michael Firko, the head of biotechnology regulation at the USDA's Animal and Plant Health Inspection Service. USDA scientists have a limited amount of plant DNA to work with, and are proceeding cautiously in testing for single-base differences or variations in the number of copies of a repeated sequence that are unique to the various strains of wheat. "It takes time," says Firko, "but we're making good progress."

► **NATURE.COM**
For more on GM crops see *Nature's* special issue:
go.nature.com/gafvrp

NATALIE BEHRING/BLOOMBERG VIA GETTY IMAGES

► in Tucson. The survey has discovered most of the small near-Earth asteroids so far. These cameras — planned for mid-2014 — would double the survey's field of view and increase the chances of catching a faint small object, says Spahr.

NASA also plans to restart the Wide-field Infrared Survey Explorer, a spacecraft that has been in hibernation since 2011, to hunt for more asteroids. Two facilities slated for Hawaii — the Pan-STARRS-2 telescope and the ATLAS telescope array — could also add to the effort.

The extra eyes are crucial because near-Earth asteroids often vanish from view within days of a sighting. On discovery, astronomers must quickly marshal follow-up observations with radar and infrared telescopes to learn more about the rock's size, shape and composition.

Engineers at NASA's Johnson Space Center in Houston, Texas, are already working on several mission designs for the robotic capture craft, which would need to grapple with an asteroid weighing around 400 tonnes — about the mass of the International Space Station (ISS). The xenon-ion thrusters used to get to the asteroid, and the inflatable rip-resistant bag used to capture it, are straightforward components. The bigger issue is finding a target and building the spacecraft in time for launch, says Andrew Thomas, an astronaut at Johnson who works on the project. "It's pretty brazen, if you think about it," he says.

A 2017 launch would allow the spacecraft to drag the target asteroid to a position near the Moon by 2021. In theory, astronauts could then visit the asteroid, peel back or cut the bag and take rock samples back to Earth.

NASA expects the mission to cost between US\$1 billion and \$2.6 billion. Agency officials are careful to emphasize that the goal is not science, but rather to give astronauts practice in going into deep space, where they must contend with harsher radiation environments than on missions to the ISS. "If our goal is to get to Mars in the 2030 time frame, we're going to have to be aggressive with this upfront stuff," says William Gerstenmaier, NASA's associate administrator for human exploration and operations.

It is not yet clear whether Obama's plan will be paid for by Congress, which has been locked in a battle with the administration over agency goals ever since NASA gave up former President George W. Bush's vision of returning to the Moon (see *Nature* **492**, 161–162; 2012). A bill proposed by a congressional subcommittee on space in the Republican-dominated House of Representatives would cancel all funding for the asteroid initiative. ■



Wheat in a test field at Oregon State University where scientists are helping to identify a rogue GM crop.

BIOTECHNOLOGY

Hunt for mystery GM wheat heats up

Investigators hope to track origins of the transgenic crop.

BY HEIDI LEDFORD

It has been nearly three months since an Oregon farmer discovered unapproved transgenic wheat in a commercial wheat field, triggering bans on imports of US wheat into Japan and South Korea. The harvest season has now begun, and with the contamination proving to be an isolated event, imports into South Korea have resumed.

But as an army of combines marches across the wheat fields of eastern Oregon, the mystery of the transgenic intruders is fresh in the minds of investigators at the US Department of Agriculture (USDA), who are trying to trace the plants' provenance to a particular research plot. Those close to the investigation say that the identity of the variety in question could emerge in the coming weeks, providing a much-needed breakthrough in this agricultural whodunnit. It could point to the cause of the release, which some have suggested could be activist sabotage. "We may never know who actually released it," says James Moyer, director of the Agricultural Research Center at Washington State University in Pullman. "But if they know the genotype of those plants, they will be able to narrow it down quite a bit."

Within a month of the discovery in May, USDA scientists had traced the origin of the plants to a line of herbicide-resistant 'Roundup

Ready' wheat called MON71800, developed by the agricultural company Monsanto, based in St Louis, Missouri. Monsanto killed the project in 2005 over farmers' worries that overseas customers would not buy US wheat if it contained transgenic varieties. No GM wheat has yet been approved to be grown commercially in the United States. The company says that all seed from the field trials — conducted on more than 400 hectares in 16 states (see 'Sifting for GM wheat') — was accounted for and either secured or destroyed.

Monsanto had shipped MON71800 seed to breeders around the country for crossing with commercial varieties optimized for each region's climate, day length and disease profile. Now, the USDA investigators are sifting through hundreds of markers to try to match the genetic signature of the contaminant Oregon wheat with one of the varieties from the 256 field tests registered with the USDA. An origin near Oregon could hint at an accidental escape; seeds from farther afield could mean that someone had intentionally saved seed and released it.

But pinning down the variety is difficult, says Michael Firko, the head of biotechnology regulation at the USDA's Animal and Plant Health Inspection Service. USDA scientists have a limited amount of plant DNA to work with, and are proceeding cautiously in testing for single-base differences or variations in the number of copies of a repeated sequence that are unique to the various strains of wheat. "It takes time," says Firko, "but we're making good progress."

► **NATURE.COM**
For more on GM crops see *Nature's* special issue:
go.nature.com/gafvrp

NATALIE BEHRING/BLOOMBERG VIA GETTY IMAGES

SOURCE: USDA, MONSANTO

Monsanto has already made clear its favoured explanation for the contamination: sabotage. “There are folks who don’t like biotechnology and who would use this as an opportunity to create problems,” the company’s chief technology officer Robert Fraley told reporters last month. Activists opposed to genetically modified crops are best known for destroying the plants rather than sowing them, but Fraley argues that those who illegally enter fields to demolish crops could also break into experimental plots to collect seed.

That hypothesis has little support among plant scientists contacted by *Nature*. “I suppose it’s possible, but it’s a very small possibility,” says Norman Ellstrand, a plant biologist at the University of California, Riverside. He says that any saboteur would have been taking a gamble that the GM wheat would be found. It only came to light in Oregon because the farmer had sprayed a non-GM wheat field with glyphosate (the herbicide to which the GM wheat is resistant) in preparation for a new crop, noticed a few remaining wheat plants and notified others of the discovery.

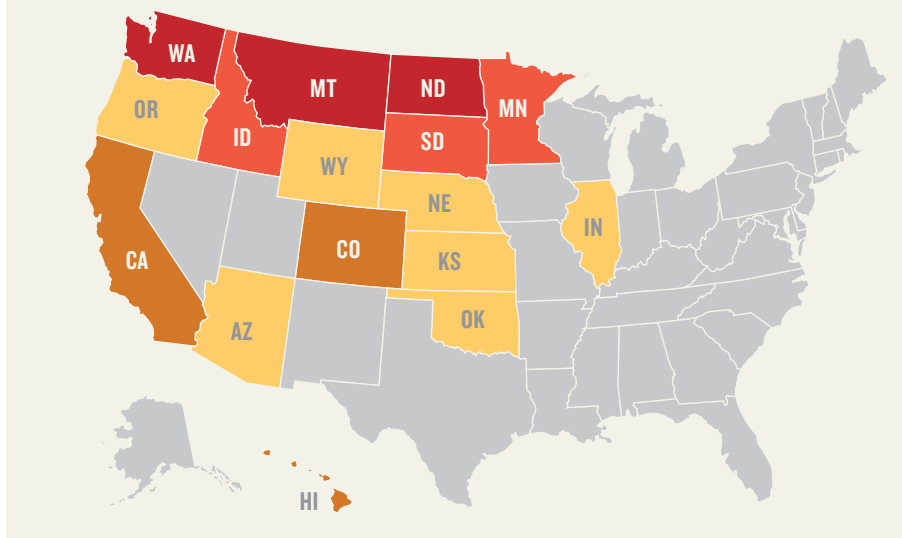
Fraley argues that the distribution of the contaminant plants suggests that a human hand cast them there. They were found in localized patches in only one of two wheat fields that had been planted with the same non-GM seed. But Robert Zemetra, a wheat breeder at Oregon State University (OSU) in Corvallis, says there could be other explanations for such a distribution. If, for example, the contaminant was a spring wheat plant in a winter wheat field, the transgenic wheat would flower and drop most of its seeds before the rest of the crop was harvested. Those seeds

SIFTING FOR GM WHEAT

Between 1997 and 2005, Monsanto conducted 256 field trials of its herbicide-resistant wheat in 16 states. Genetic testing could help determine which of these GM varieties wound up in an Oregon wheat field.

Total number of field trials

2–10 11–20
21–30 >30



would fall straight down, generating a clump of herbicide-resistant offspring.

No explanation is completely satisfying, acknowledges Rene Van Acker, a weed scientist at the University of Guelph in Ontario, Canada. Van Acker and Zemetra carried out separate field trials of the wheat over a decade ago, and both say that Monsanto kept a close watch over the experiments. “We had to account for pretty much every seed in and every seed out, down to the gram,” recalls Van Acker. Zemetra not-so-fondly remembers hours spent disassembling harvesters and cleaning each part with an

air-compressor, removing seeds with tweezers when necessary.

But Carol Mallory-Smith, the OSU weed scientist who first tested the Oregon plants three months ago, wouldn’t be surprised if one of the field-test seeds had escaped. She has found transgenic crops in stranger places. In 2009, for instance, she found transgenic sugar-beet seedlings in a bag of soil sold to gardeners. “There are so many places in the system where errors can be made,” she says. “Once we release these genes into the field, we should just assume that they are going to stay in the environment.” ■

MEDICINE

NIH gambles on recycled drugs

Early success could bolster congressional support for agency’s translational science centre.

BY MEREDITH WADMAN

On Monday, Stephen Strittmatter began recruiting for a clinical trial to test a treatment for Alzheimer’s disease. The study is also part of a bigger test for the US National Institutes of Health (NIH) and its fledgling translational research centre.

Strittmatter’s team is one of nine that won funding last month from the NIH’s National Center for Advancing Translational Sciences (NCATS) in Bethesda, Maryland, to see whether abandoned drugs can be aimed at new targets. Strittmatter, a neurobiologist at Yale University in New Haven, Connecticut, hopes that a failed cancer drug called saracatinib can

block an enzyme implicated in Alzheimer’s.

The money is not much — US\$12.7 million split between the nine awards this year — but the symbolism is potent. NCATS, which launched in December 2011, has promised to find ways around obstacles that stymie the pharmaceutical industry. Legislators have been sceptical about the need for such a programme, but if the drug-resurrection initiative succeeds, it may win over critics in Congress, for whom repurposing drugs has appeal in an age of fiscal austerity.

Agency officials downplay the stakes, arguing that the programme’s chief aim is to establish new partnerships between industry and academia. Critics say that its scientific goal

may be a tall order. “Others passed [on these drugs] because the scientific odds of success were low,” says Scott Gottlieb, a resident fellow at the conservative American Enterprise Institute in Washington DC, and a former deputy commissioner of the Food and Drug Administration. “The odds that NIH has some early wins are equally low.”

The projects are tackling conditions including alcoholism, schizophrenia and calcified aortic valves (See ‘Old drugs, new tricks’). But none targets an illness of more growing public health concern than Alzheimer’s, which is expected to afflict up to 16 million people in the United States by 2050, and will cost the country around \$200 billion in 2013. ▶

SOURCE: USDA, MONSANTO

Monsanto has already made clear its favoured explanation for the contamination: sabotage. “There are folks who don’t like biotechnology and who would use this as an opportunity to create problems,” the company’s chief technology officer Robert Fraley told reporters last month. Activists opposed to genetically modified crops are best known for destroying the plants rather than sowing them, but Fraley argues that those who illegally enter fields to demolish crops could also break into experimental plots to collect seed.

That hypothesis has little support among plant scientists contacted by *Nature*. “I suppose it’s possible, but it’s a very small possibility,” says Norman Ellstrand, a plant biologist at the University of California, Riverside. He says that any saboteur would have been taking a gamble that the GM wheat would be found. It only came to light in Oregon because the farmer had sprayed a non-GM wheat field with glyphosate (the herbicide to which the GM wheat is resistant) in preparation for a new crop, noticed a few remaining wheat plants and notified others of the discovery.

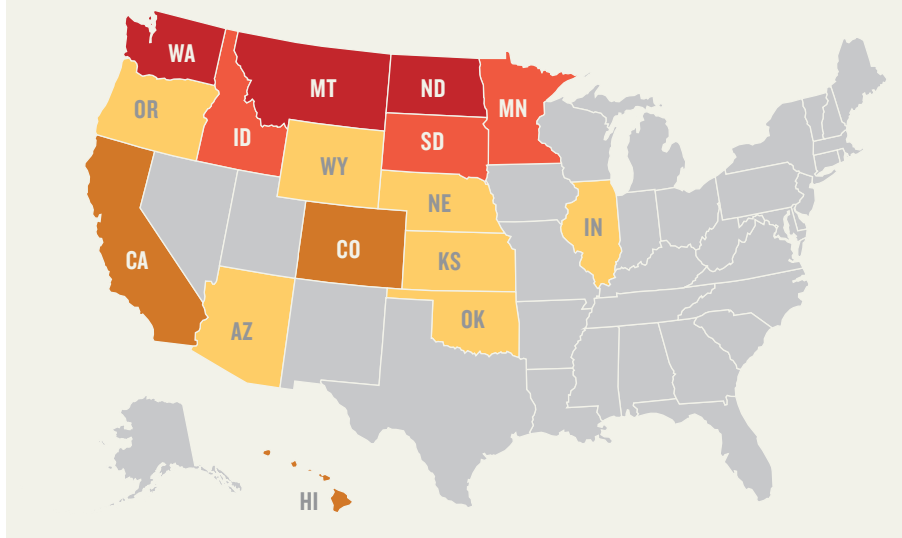
Fraley argues that the distribution of the contaminant plants suggests that a human hand cast them there. They were found in localized patches in only one of two wheat fields that had been planted with the same non-GM seed. But Robert Zemetra, a wheat breeder at Oregon State University (OSU) in Corvallis, says there could be other explanations for such a distribution. If, for example, the contaminant was a spring wheat plant in a winter wheat field, the transgenic wheat would flower and drop most of its seeds before the rest of the crop was harvested. Those seeds

SIFTING FOR GM WHEAT

Between 1997 and 2005, Monsanto conducted 256 field trials of its herbicide-resistant wheat in 16 states. Genetic testing could help determine which of these GM varieties wound up in an Oregon wheat field.

Total number of field trials

2–10 11–20
21–30 >30



would fall straight down, generating a clump of herbicide-resistant offspring.

No explanation is completely satisfying, acknowledges Rene Van Acker, a weed scientist at the University of Guelph in Ontario, Canada. Van Acker and Zemetra carried out separate field trials of the wheat over a decade ago, and both say that Monsanto kept a close watch over the experiments. “We had to account for pretty much every seed in and every seed out, down to the gram,” recalls Van Acker. Zemetra not-so-fondly remembers hours spent disassembling harvesters and cleaning each part with an

air-compressor, removing seeds with tweezers when necessary.

But Carol Mallory-Smith, the OSU weed scientist who first tested the Oregon plants three months ago, wouldn’t be surprised if one of the field-test seeds had escaped. She has found transgenic crops in stranger places. In 2009, for instance, she found transgenic sugar-beet seedlings in a bag of soil sold to gardeners. “There are so many places in the system where errors can be made,” she says. “Once we release these genes into the field, we should just assume that they are going to stay in the environment.” ■

MEDICINE

NIH gambles on recycled drugs

Early success could bolster congressional support for agency’s translational science centre.

BY MEREDITH WADMAN

On Monday, Stephen Strittmatter began recruiting for a clinical trial to test a treatment for Alzheimer’s disease. The study is also part of a bigger test for the US National Institutes of Health (NIH) and its fledgling translational research centre.

Strittmatter’s team is one of nine that won funding last month from the NIH’s National Center for Advancing Translational Sciences (NCATS) in Bethesda, Maryland, to see whether abandoned drugs can be aimed at new targets. Strittmatter, a neurobiologist at Yale University in New Haven, Connecticut, hopes that a failed cancer drug called saracatinib can

block an enzyme implicated in Alzheimer’s.

The money is not much — US\$12.7 million split between the nine awards this year — but the symbolism is potent. NCATS, which launched in December 2011, has promised to find ways around obstacles that stymie the pharmaceutical industry. Legislators have been sceptical about the need for such a programme, but if the drug-resurrection initiative succeeds, it may win over critics in Congress, for whom repurposing drugs has appeal in an age of fiscal austerity.

Agency officials downplay the stakes, arguing that the programme’s chief aim is to establish new partnerships between industry and academia. Critics say that its scientific goal

may be a tall order. “Others passed [on these drugs] because the scientific odds of success were low,” says Scott Gottlieb, a resident fellow at the conservative American Enterprise Institute in Washington DC, and a former deputy commissioner of the Food and Drug Administration. “The odds that NIH has some early wins are equally low.”

The projects are tackling conditions including alcoholism, schizophrenia and calcified aortic valves (See ‘Old drugs, new tricks’). But none targets an illness of more growing public health concern than Alzheimer’s, which is expected to afflict up to 16 million people in the United States by 2050, and will cost the country around \$200 billion in 2013. ▶

OLD DRUGS, NEW TRICKS

The US National Institutes of Health has awarded nine grants to find uses for eight drugs abandoned by pharmaceutical firms.

Drug name	Original target(s)	New target(s)
PF-03463275	Schizophrenia	Schizophrenia (new mechanism)
AZD0530 (saracatinib)	Cancer	Alzheimer's disease; lymphangioleiomyomatosis
JNJ-39393406	Schizophrenia	Smoking cessation
ZD4054 (zibotentan)	Cancer; peripheral artery hypertension	Peripheral artery disease
PF-05190457	Type II diabetes	Alcoholism
LY500307	Prostate enlargement; urinary tract symptoms	Schizophrenia
Undisclosed (made by Sanofi)	Undisclosed	Calcific aortic valve stenosis
Undisclosed (made by Sanofi)	Undisclosed	Duchenne muscular dystrophy

Source: National Institutes of Health

► Saracatinib inhibits the Src family kinases (SFKs), enzymes that are commonly activated in cancer cells, and was first developed by London-based pharmaceutical company AstraZeneca. But the drug proved only marginally effective against cancer, and the company abandoned it — after spending millions of dollars to develop it through early human trials that proved that it was safe. With that work already done, Strittmatter's group will be able to move the drug quickly into testing in people with early-stage Alzheimer's disease.

The team plans to begin a 24-person safety and dosing trial in August. If the results are good, NCATS will fund the effort for two more years, during which the scientists will launch a double-blind, randomized, placebo-controlled

trial with 159 participants. Over a year, the team will measure declines in glucose metabolism — a marker for progression of Alzheimer's disease — in key brain regions, hoping to find that they have slowed.

"For this to happen in years instead of decades is only possible because of this programme," says Strittmatter. The NIH worked with eight pharmaceutical companies to develop a model agreement for the firms and academic investigators; it specifies how intellectual property will be apportioned and gives the companies first right of refusal to license any discoveries.

Now NCATS must wait for results. Using saracatinib to treat Alzheimer's is "a reasonable idea", says Tony Hunter, director of the Salk Institute Cancer Center in La Jolla, California.

The Yale researchers hope that the drug will slow or halt progression of the disease by blocking the activity of Fyn kinase, an SFK that has been shown to induce damage to brain synapses in mouse models of Alzheimer's (J. Chin *et al. J. Neurosci.* **24**, 4692–4697; 2004). But Hunter warns that there may be side effects. SFKs have many functions, so if high doses of saracatinib are needed to inhibit Fyn kinase action in the brain, they could do damage elsewhere.

Strittmatter says that the main side effect of concern to him is reduced white-blood-cell counts, which could increase risk of infection. However, in saracatinib cancer trials, that side effect occurred at higher doses than he intends to use.

NCATS is barred by law from funding trials beyond early phase II, so further money will have to come from another backer, such as AstraZeneca. For that, observers say, Strittmatter will need startlingly good results.

"This is going to be hard for them to do," says John LaMattina, a senior partner at PureTech Ventures, a life-sciences venture-capital company in Boston, Massachusetts, and former president of research and development at the drug-maker Pfizer in New York.

But fans of the NIH programme say that it is not reasonable to expect quick success, given the time-consuming nature of drug development. "The commitment to creative repurposing is wise and timely," says Ann Bonham, chief scientific officer at the Association of American Medical Colleges in Washington DC. "Over time, even with a relatively modest success rate, the potential benefit is worth the investments." ■

ANIMAL DISEASE

Rinderpest research restarts

As moratorium lifts, oversight is put in place to assess studies on eradicated cattle virus.

BY DECLAN BUTLER

Research is set to resume on the rinderpest virus, the cause of a deadly cattle disease that was declared eradicated in 2011 and has been off limits for study ever since. The moratorium — part of efforts to guard against accidental or intentional release of virus that could reintroduce the disease — was lifted on 10 July and replaced by a new international oversight system for such research.

In its heyday, the disease — the only one other than smallpox to be eradicated from nature — killed hundreds of millions of cattle, mainly in Europe, Asia and Africa, often leaving famine in its wake. Under the new oversight system, run by the Food and Agriculture

Organization of the United Nations (FAO) in Rome and the Paris-based World Organisation for Animal Health (OIE), the risks and benefits of research proposals will be assessed by a joint advisory committee, and then the FAO and the OIE will decide on approvals. Eligible research must show potential for substantial practical or scientific benefits and be conducted under stringent biosafety and biosecurity conditions.

The first project that has garnered approval will test whether vaccines developed against a closely related virus — *peste des petits ruminants* (PPR), which causes disease in sheep and goats — might also protect cattle against rinderpest. Led by Michael Baron,

a rinderpest researcher at the Pirbright Institute in Pirbright, UK, the project, if successful, would eliminate the need to retain stocks of live-attenuated rinderpest vaccine. That would contribute to the goal of reducing the number of labs worldwide holding rinderpest material, thus decreasing the risk of reintroduction.

Some 55 labs in 35 countries still hold some kind of rinderpest virus, according to a 2011 survey published in January 2013 in the journal *Emerging Infectious Diseases*: 37% of them in Asia, 29% in Africa and 26% in Europe (G. Fournié *et al. Emerging Infect. Dis.* <http://doi.org/m7w>; 2013). The identities of the labs remain confidential. The most dangerous stocks are of live field strains of virus, estimated to be kept in at least 16 labs in 14 countries, and samples of blood and tissues from infected herds,

► NATURE.COM

For more on
rinderpest, see:
go.nature.com/nait5m

OLD DRUGS, NEW TRICKS

The US National Institutes of Health has awarded nine grants to find uses for eight drugs abandoned by pharmaceutical firms.

Drug name	Original target(s)	New target(s)
PF-03463275	Schizophrenia	Schizophrenia (new mechanism)
AZD0530 (saracatinib)	Cancer	Alzheimer's disease; lymphangioleiomyomatosis
JNJ-39393406	Schizophrenia	Smoking cessation
ZD4054 (zibotentan)	Cancer; peripheral artery hypertension	Peripheral artery disease
PF-05190457	Type II diabetes	Alcoholism
LY500307	Prostate enlargement; urinary tract symptoms	Schizophrenia
Undisclosed (made by Sanofi)	Undisclosed	Calcific aortic valve stenosis
Undisclosed (made by Sanofi)	Undisclosed	Duchenne muscular dystrophy

Source: National Institutes of Health

► Saracatinib inhibits the Src family kinases (SFKs), enzymes that are commonly activated in cancer cells, and was first developed by London-based pharmaceutical company AstraZeneca. But the drug proved only marginally effective against cancer, and the company abandoned it — after spending millions of dollars to develop it through early human trials that proved that it was safe. With that work already done, Strittmatter's group will be able to move the drug quickly into testing in people with early-stage Alzheimer's disease.

The team plans to begin a 24-person safety and dosing trial in August. If the results are good, NCATS will fund the effort for two more years, during which the scientists will launch a double-blind, randomized, placebo-controlled

trial with 159 participants. Over a year, the team will measure declines in glucose metabolism — a marker for progression of Alzheimer's disease — in key brain regions, hoping to find that they have slowed.

"For this to happen in years instead of decades is only possible because of this programme," says Strittmatter. The NIH worked with eight pharmaceutical companies to develop a model agreement for the firms and academic investigators; it specifies how intellectual property will be apportioned and gives the companies first right of refusal to license any discoveries.

Now NCATS must wait for results. Using saracatinib to treat Alzheimer's is "a reasonable idea", says Tony Hunter, director of the Salk Institute Cancer Center in La Jolla, California.

The Yale researchers hope that the drug will slow or halt progression of the disease by blocking the activity of Fyn kinase, an SFK that has been shown to induce damage to brain synapses in mouse models of Alzheimer's (J. Chin *et al. J. Neurosci.* **24**, 4692–4697; 2004). But Hunter warns that there may be side effects. SFKs have many functions, so if high doses of saracatinib are needed to inhibit Fyn kinase action in the brain, they could do damage elsewhere.

Strittmatter says that the main side effect of concern to him is reduced white-blood-cell counts, which could increase risk of infection. However, in saracatinib cancer trials, that side effect occurred at higher doses than he intends to use.

NCATS is barred by law from funding trials beyond early phase II, so further money will have to come from another backer, such as AstraZeneca. For that, observers say, Strittmatter will need startlingly good results.

"This is going to be hard for them to do," says John LaMattina, a senior partner at PureTech Ventures, a life-sciences venture-capital company in Boston, Massachusetts, and former president of research and development at the drug-maker Pfizer in New York.

But fans of the NIH programme say that it is not reasonable to expect quick success, given the time-consuming nature of drug development. "The commitment to creative repurposing is wise and timely," says Ann Bonham, chief scientific officer at the Association of American Medical Colleges in Washington DC. "Over time, even with a relatively modest success rate, the potential benefit is worth the investments." ■

ANIMAL DISEASE

Rinderpest research restarts

As moratorium lifts, oversight is put in place to assess studies on eradicated cattle virus.

BY DECLAN BUTLER

Research is set to resume on the rinderpest virus, the cause of a deadly cattle disease that was declared eradicated in 2011 and has been off limits for study ever since. The moratorium — part of efforts to guard against accidental or intentional release of virus that could reintroduce the disease — was lifted on 10 July and replaced by a new international oversight system for such research.

In its heyday, the disease — the only one other than smallpox to be eradicated from nature — killed hundreds of millions of cattle, mainly in Europe, Asia and Africa, often leaving famine in its wake. Under the new oversight system, run by the Food and Agriculture

Organization of the United Nations (FAO) in Rome and the Paris-based World Organisation for Animal Health (OIE), the risks and benefits of research proposals will be assessed by a joint advisory committee, and then the FAO and the OIE will decide on approvals. Eligible research must show potential for substantial practical or scientific benefits and be conducted under stringent biosafety and biosecurity conditions.

The first project that has garnered approval will test whether vaccines developed against a closely related virus — *peste des petits ruminants* (PPR), which causes disease in sheep and goats — might also protect cattle against rinderpest. Led by Michael Baron,

a rinderpest researcher at the Pirbright Institute in Pirbright, UK, the project, if successful, would eliminate the need to retain stocks of live-attenuated rinderpest vaccine. That would contribute to the goal of reducing the number of labs worldwide holding rinderpest material, thus decreasing the risk of reintroduction.

Some 55 labs in 35 countries still hold some kind of rinderpest virus, according to a 2011 survey published in January 2013 in the journal *Emerging Infectious Diseases*: 37% of them in Asia, 29% in Africa and 26% in Europe (G. Fournié *et al. Emerging Infect. Dis.* <http://doi.org/m7w>; 2013). The identities of the labs remain confidential. The most dangerous stocks are of live field strains of virus, estimated to be kept in at least 16 labs in 14 countries, and samples of blood and tissues from infected herds,

► **NATURE.COM**

For more on
rinderpest, see:
go.nature.com/nait5m

kept in at least 10 labs in 10 countries. Stocks of live-attenuated vaccine, currently held in at least 53 labs in 34 countries, are deemed less problematic, although some could, in theory, revert to disease-causing forms.

The FAO and the OIE hope to eventually reduce the number of sites holding live wild viruses to a handful of officially designated labs, ideally located outside regions where accidental releases could have devastating consequences, says David Ulaeto, a virologist and member of the joint advisory committee. Conversely, the agencies plan to centralize stocks of vaccines in a few high-containment repositories in regions at highest risk of disease, so that they can be deployed within hours of any confirmed recurrence of rinderpest. No siting decisions have been made, but one might imagine a repository in Africa, one or two in Asia and one in Europe, says Juan Lubroth, the FAO's chief veterinary officer.

The process of destroying virus or shipping it to centres with high biosafety levels must be done in a way that does not risk its release, says Ulaeto. The FAO and the OIE are working on high-security protocols for shipping the virus and ways to ensure that autoclaves in labs holding it are certified to function at levels guaranteed to provide a 100% kill.

Many countries are reluctant to give up their vaccine stocks in case the disease should reappear and threaten their food supply. They worry about becoming dependent on the willingness of the international community to swiftly provide them with needed vaccines. "The challenge is political," says Bernard Vallat, director general of the OIE. He says that the FAO and the OIE are drafting agreements and international contingency plans that should help reassure countries that swift help would be forthcoming and that they would have guaranteed access to vaccine from FAO-OIE repositories.

Vallat notes that if Baron proves that PPR vaccines can protect cattle against rinderpest, it would provide an elegant way around such political issues: there would no longer be any need to hold onto rinderpest vaccines. Baron says that he hopes to start the vaccine-challenge trials next spring and complete them by the end of 2014.

Additional potentially promising research areas include other improved vaccines, diagnostics and perhaps disease pathology, says Lubroth. He stresses, however, that the advisory committee will not be prescriptive but open to considering any research ideas put forward by scientists. ■

Many countries are reluctant to give up their vaccine stocks.

RADIATION BIOLOGY

Fukushima offers real-time ecolab

But ecologists say they need more funding.

BY EWEN CALLAWAY

Hours after a magnitude-9 earthquake struck off Japan's eastern coast in March 2011 and triggered the Fukushima nuclear disaster, Marta Wayne e-mailed colleagues in Japan — first to check on their safety, and later to make plans.

The 1986 meltdown of a nuclear reactor at Chernobyl in Ukraine had been a missed opportunity for researchers to gather data on the ecological effects of low-level radiation, she says. Independent scientists didn't gain access to the area for a decade. This time, "I thought immediately that it was important to seize the moment, and study and get data on what the actual outcomes of such a disaster could be", says Wayne, a population geneticist at the University of Florida, Gainesville.

Last week, Wayne and other biologists studying Fukushima and Chernobyl came together at the annual meeting of the Society for Molecular Biology and Evolution in Chicago, Illinois, to report on what they'd learned so far — and what studies they feel are needed for the future. They believe their work on the effects of low-level radiation on animals such as butterflies and sparrows is relevant to understanding the impact of low-level radiation on humans, with implications for appropriate government responses to radiation releases.

The effects of such exposure in humans are poorly understood, says David Brenner, director of the Center for Radiological Research at Columbia University in New York. In an 18 March letter to John Holdren, the US president's chief science adviser, he and his colleagues called for a comprehensive research strategy on the problem. "We're stuck in a dilemma about having to make policy decisions based on nothing more than guesses," he says.

Brenner adds that the risks — mainly of cancer — are small. A March 2013 report by the World Health Organization in Geneva, Switzerland, identified hotspots in Fukushima prefecture where it is predicted that children may experience a slightly increased overall risk of some rare cancers, such as those affecting the thyroid. But most human epidemiological studies are not big enough to pick up small increases in the prevalence of rare conditions.



Zizeeria maha with abnormal wings.

Scientists such as Wayne think they can fill in some of the knowledge gaps by studying other species, if they can secure the necessary funding. That has proved enormously difficult in a world where data on the effects of radiation are the subject of heated debate in arguments over the use of nuclear energy.

What Fukushima data do exist are sporadic — and contested. One research flurry concerns butterflies. Joji Otaki, an ecologist at the University of the Ryukyus in Nishihara, Japan, has for more than a decade studied the wing-spot patterns and other traits in a Japanese species, *Zizeeria maha*. "I never dreamed of using it for a nuclear accident," says Otaki, who presented his work at the Chicago meeting. But after the Fukushima meltdown, two of his graduate students convinced him to screen for abnormalities in the butterfly as an environmental indicator of radiation's effects.

The team went to Fukushima in May 2011, two months after the earthquake, when the butterflies emerge from their chrysalises, and again in September 2011. They collected butterflies from sites ranging from 20 to 225 km from the reactor. Insects collected in May showed few problems, but their lab-reared offspring had many abnormalities, such as misshapen wings and aberrant eyespots, and many died as pupae (A. Hiyama *et al. Sci. Rep.* 2, 570; 2012). Among the September-collected butterflies, more than half of the progeny showed such defects.

Otaki's team also exposed butterflies to radiation doses in the lab akin to those that butterflies near Fukushima might have received. The offspring developed the same problems. "You can come up with alternative explanations, but I think the hypothesis that radiation caused death and abnormalities is the most reasonable," Otaki says. ▶

➔ **NATURE.COM**
For more on the Fukushima disaster see:
go.nature.com/ulsz2n

CHIYO NOHARA & JOJI M. OTAKI

kept in at least 10 labs in 10 countries. Stocks of live-attenuated vaccine, currently held in at least 53 labs in 34 countries, are deemed less problematic, although some could, in theory, revert to disease-causing forms.

The FAO and the OIE hope to eventually reduce the number of sites holding live wild viruses to a handful of officially designated labs, ideally located outside regions where accidental releases could have devastating consequences, says David Ulaeto, a virologist and member of the joint advisory committee. Conversely, the agencies plan to centralize stocks of vaccines in a few high-containment repositories in regions at highest risk of disease, so that they can be deployed within hours of any confirmed recurrence of rinderpest. No siting decisions have been made, but one might imagine a repository in Africa, one or two in Asia and one in Europe, says Juan Lubroth, the FAO's chief veterinary officer.

The process of destroying virus or shipping it to centres with high biosafety levels must be done in a way that does not risk its release, says Ulaeto. The FAO and the OIE are working on high-security protocols for shipping the virus and ways to ensure that autoclaves in labs holding it are certified to function at levels guaranteed to provide a 100% kill.

Many countries are reluctant to give up their vaccine stocks in case the disease should reappear and threaten their food supply. They worry about becoming dependent on the willingness of the international community to swiftly provide them with needed vaccines. "The challenge is political," says Bernard Vallat, director general of the OIE. He says that the FAO and the OIE are drafting agreements and international contingency plans that should help reassure countries that swift help would be forthcoming and that they would have guaranteed access to vaccine from FAO-OIE repositories.

Vallat notes that if Baron proves that PPR vaccines can protect cattle against rinderpest, it would provide an elegant way around such political issues: there would no longer be any need to hold onto rinderpest vaccines. Baron says that he hopes to start the vaccine-challenge trials next spring and complete them by the end of 2014.

Additional potentially promising research areas include other improved vaccines, diagnostics and perhaps disease pathology, says Lubroth. He stresses, however, that the advisory committee will not be prescriptive but open to considering any research ideas put forward by scientists. ■

Many countries are reluctant to give up their vaccine stocks.

RADIATION BIOLOGY

Fukushima offers real-time ecolab

But ecologists say they need more funding.

BY EWEN CALLAWAY

Hours after a magnitude-9 earthquake struck off Japan's eastern coast in March 2011 and triggered the Fukushima nuclear disaster, Marta Wayne e-mailed colleagues in Japan — first to check on their safety, and later to make plans.

The 1986 meltdown of a nuclear reactor at Chernobyl in Ukraine had been a missed opportunity for researchers to gather data on the ecological effects of low-level radiation, she says. Independent scientists didn't gain access to the area for a decade. This time, "I thought immediately that it was important to seize the moment, and study and get data on what the actual outcomes of such a disaster could be", says Wayne, a population geneticist at the University of Florida, Gainesville.

Last week, Wayne and other biologists studying Fukushima and Chernobyl came together at the annual meeting of the Society for Molecular Biology and Evolution in Chicago, Illinois, to report on what they'd learned so far — and what studies they feel are needed for the future. They believe their work on the effects of low-level radiation on animals such as butterflies and sparrows is relevant to understanding the impact of low-level radiation on humans, with implications for appropriate government responses to radiation releases.

The effects of such exposure in humans are poorly understood, says David Brenner, director of the Center for Radiological Research at Columbia University in New York. In an 18 March letter to John Holdren, the US president's chief science adviser, he and his colleagues called for a comprehensive research strategy on the problem. "We're stuck in a dilemma about having to make policy decisions based on nothing more than guesses," he says.

Brenner adds that the risks — mainly of cancer — are small. A March 2013 report by the World Health Organization in Geneva, Switzerland, identified hotspots in Fukushima prefecture where it is predicted that children may experience a slightly increased overall risk of some rare cancers, such as those affecting the thyroid. But most human epidemiological studies are not big enough to pick up small increases in the prevalence of rare conditions.



Zizeeria maha with abnormal wings.

Scientists such as Wayne think they can fill in some of the knowledge gaps by studying other species, if they can secure the necessary funding. That has proved enormously difficult in a world where data on the effects of radiation are the subject of heated debate in arguments over the use of nuclear energy.

What Fukushima data do exist are sporadic — and contested. One research flurry concerns butterflies. Joji Otaki, an ecologist at the University of the Ryukyus in Nishihara, Japan, has for more than a decade studied the wing-spot patterns and other traits in a Japanese species, *Zizeeria maha*. "I never dreamed of using it for a nuclear accident," says Otaki, who presented his work at the Chicago meeting. But after the Fukushima meltdown, two of his graduate students convinced him to screen for abnormalities in the butterfly as an environmental indicator of radiation's effects.

The team went to Fukushima in May 2011, two months after the earthquake, when the butterflies emerge from their chrysalises, and again in September 2011. They collected butterflies from sites ranging from 20 to 225 km from the reactor. Insects collected in May showed few problems, but their lab-reared offspring had many abnormalities, such as misshapen wings and aberrant eyespots, and many died as pupae (A. Hiyama *et al. Sci. Rep.* 2, 570; 2012). Among the September-collected butterflies, more than half of the progeny showed such defects.

Otaki's team also exposed butterflies to radiation doses in the lab akin to those that butterflies near Fukushima might have received. The offspring developed the same problems. "You can come up with alternative explanations, but I think the hypothesis that radiation caused death and abnormalities is the most reasonable," Otaki says. ▶

➔ **NATURE.COM**
For more on the Fukushima disaster see:
go.nature.com/ulsz2n

CHIYO NOHARA & JOJI M. OTAKI

► Tim Mousseau, an evolutionary geneticist at the University of South Carolina in Columbia, says that more studies like these are sorely needed. He is heading to Fukushima this week to begin his third season of field work since the meltdown, tracking birds, insects and other small animals. His team saw die-offs in some insects and declining numbers of some bird populations after one season's work (A. P. Möller *et al. Environ. Pollut.* **164**, 36–39; 2012). He hopes soon to publish three years of observations.

For funding, Otaki says he has had to turn mostly to private foundations. "I think maybe this is a very touchy issue, politically," he says. Mousseau has received money from a German biotechnology company and is now working with researchers supported by the Finnish government. But he says that US government grants are difficult to secure. The Department of Energy has largely stopped funding its research programme in low-dose exposure, and the National Science Foundation and the National Institutes of Health have awarded few grants on the topic. "The only people who seem to be doing any research are adventurous, opportunistic and independent," Mousseau says. "They have some flexibility in what they do and are just doing it on their own without official support."

Other scientists take issue with the reports of ecological harms from Fukushima. They say that Otaki's research is flawed, because wing shape and other butterfly traits vary naturally with geography. "This study's sensational claims should not be used to scare the local population into the erroneous conclusion that their exposures to these relatively low environmental radiation doses put them at significant health risk," Timothy Jorgensen, a molecular radiation biologist at Georgetown University in Washington DC, wrote in a comment on Otaki's 2012 paper. Mousseau's report of harms to birds one year after Fukushima has been criticized for including only one sampling period and lacking baseline data.

Richard Wakeford, an epidemiologist at the University of Manchester, UK, thinks that ecological research on the Fukushima disaster's effects will prove as confounding as efforts to detect health effects in humans exposed to low doses of radiation. Many ecosystems and their species are altered after human evacuations in ways that have nothing to do with radiation, he says.

Wayne says post-Fukushima research needs more support to boost its quality. She and her colleagues are drafting a white paper to establish better standards for collecting, analysing and sharing data. "We don't want disasters to happen so we can collect more data," she says. "But as it has happened, we should learn from it." ■

SPACE SCIENCE

Lasers boost space communications

Optical systems set to handle planetary science's big data.

BY DEVIN POWELL

Before NASA even existed, science-fiction writer Arthur C. Clarke in 1945 imagined spacecraft that could send messages back to Earth using beams of light. After decades of setbacks and dead ends, the technology to do this is finally coming of age.

Two spacecraft set for launch in the coming weeks will carry lasers that allow data to be transferred faster than ever before. One, scheduled for take-off on 5 September, is NASA's Lunar Atmosphere and Dust Environment Explorer (LADEE), a mission that will beam video and scientific data from the Moon. The other, a European Space Agency (ESA) project called Alphasat, is due to launch on 25 July, and will be the first optical satellite to collect large amounts of scientific data from other satellites.

"This is a big step forward," says Hamid Hemmati, a specialist in optical communications at NASA's Jet Propulsion Laboratory in Pasadena, California. "Europe is going beyond demonstrations for the first time and making operational use of the technology."

These lasers could provide bigger pipes for a coming flood of space information. New Earth-observation satellites promise to deliver petabytes of data every year. Missions such as the Mars Reconnaissance Orbiter (MRO) already have constraints on the volume of data they can send back because of fluctuations in download rates tied to a spacecraft's varying distance from Earth. "Right now, we're really far from Earth, so we can't fit as many images in our downlink," says Ingrid Daubar, who works on the MRO's HiRISE camera at the University of Arizona in Tucson. Laser data highways could ultimately allow space agencies to kit their spacecraft with more sophisticated equipment, says John Keller, deputy project scientist for NASA's Lunar Reconnaissance Orbiter (LRO). That is not yet possible, he says. "We're limited by the rate at which we can download the data."

Today's spacecraft send and receive messages using radio waves. The frequencies used are hundreds of times higher than those put out by music stations on Earth and can cram in

more information, allowing orbital broadcasts to transmit hundreds of megabits of information per second. Lasers, which operate at higher frequencies still, can reach gigabits per second (see 'Tuned in'). And unlike the radio portion of the electromagnetic spectrum, which is crowded and carefully apportioned, optical wavelengths are underused and unregulated.

Efforts to develop laser communication systems struggled for much of the twentieth century: weak lasers and problematic detectors derailed project after project. But recent advances in optics have begun to change the situation. "The technology has matured," says Frank Heine, chief scientist at Tesat-Spacecom, a company based in Backnang, Germany.

In the 1980s, Europe took advantage of improved lasers and optical detectors to begin work on its first laser communication system, the Semiconductor Laser Intersatellite Link Experiment (SILEX). Equipped with the system, the ESA satellite Artemis received 50 megabits of information per second from a French satellite in 2001 and then exchanged messages with a Japanese satellite in 2005. The project taught engineers how to stabilize and point a laser in space. But it was abandoned after its intended application — a constellation of satellites to provide Internet services — was dropped in favour of the network of fibre-optic cables now criss-crossing the globe.

Since then, Heine's team at Tesat-Spacecom has created a laser terminal for satellite-to-satellite communication, at a cost to the German Aerospace Center of €95 million (US\$124 million). The laser, amplified by modern fibre-optic technology, achieves a power of watts — compared with the tens of milliwatts reached by SILEX. In 2008, terminals mounted on two satellites transferred information at gigabits per second over a few thousand kilometres.

ESA's Alphasat will extend the range of this laser terminal to tens of thousands of kilometres once it is positioned high in geostationary orbit. Future satellites that sport laser terminals in lower orbits will be able to beam as much as 1.8 gigabits per second of information up to Alphasat, which will then relay the data to the ground using radio waves. Alphasat's geostationary orbit means that it can provide

"Laser communication becomes more advantageous the farther out you go."

► Tim Mousseau, an evolutionary geneticist at the University of South Carolina in Columbia, says that more studies like these are sorely needed. He is heading to Fukushima this week to begin his third season of field work since the meltdown, tracking birds, insects and other small animals. His team saw die-offs in some insects and declining numbers of some bird populations after one season's work (A. P. Möller *et al. Environ. Pollut.* **164**, 36–39; 2012). He hopes soon to publish three years of observations.

For funding, Otaki says he has had to turn mostly to private foundations. "I think maybe this is a very touchy issue, politically," he says. Mousseau has received money from a German biotechnology company and is now working with researchers supported by the Finnish government. But he says that US government grants are difficult to secure. The Department of Energy has largely stopped funding its research programme in low-dose exposure, and the National Science Foundation and the National Institutes of Health have awarded few grants on the topic. "The only people who seem to be doing any research are adventurous, opportunistic and independent," Mousseau says. "They have some flexibility in what they do and are just doing it on their own without official support."

Other scientists take issue with the reports of ecological harms from Fukushima. They say that Otaki's research is flawed, because wing shape and other butterfly traits vary naturally with geography. "This study's sensational claims should not be used to scare the local population into the erroneous conclusion that their exposures to these relatively low environmental radiation doses put them at significant health risk," Timothy Jorgensen, a molecular radiation biologist at Georgetown University in Washington DC, wrote in a comment on Otaki's 2012 paper. Mousseau's report of harms to birds one year after Fukushima has been criticized for including only one sampling period and lacking baseline data.

Richard Wakeford, an epidemiologist at the University of Manchester, UK, thinks that ecological research on the Fukushima disaster's effects will prove as confounding as efforts to detect health effects in humans exposed to low doses of radiation. Many ecosystems and their species are altered after human evacuations in ways that have nothing to do with radiation, he says.

Wayne says post-Fukushima research needs more support to boost its quality. She and her colleagues are drafting a white paper to establish better standards for collecting, analysing and sharing data. "We don't want disasters to happen so we can collect more data," she says. "But as it has happened, we should learn from it." ■

SPACE SCIENCE

Lasers boost space communications

Optical systems set to handle planetary science's big data.

BY DEVIN POWELL

Before NASA even existed, science-fiction writer Arthur C. Clarke in 1945 imagined spacecraft that could send messages back to Earth using beams of light. After decades of setbacks and dead ends, the technology to do this is finally coming of age.

Two spacecraft set for launch in the coming weeks will carry lasers that allow data to be transferred faster than ever before. One, scheduled for take-off on 5 September, is NASA's Lunar Atmosphere and Dust Environment Explorer (LADEE), a mission that will beam video and scientific data from the Moon. The other, a European Space Agency (ESA) project called Alphasat, is due to launch on 25 July, and will be the first optical satellite to collect large amounts of scientific data from other satellites.

"This is a big step forward," says Hamid Hemmati, a specialist in optical communications at NASA's Jet Propulsion Laboratory in Pasadena, California. "Europe is going beyond demonstrations for the first time and making operational use of the technology."

These lasers could provide bigger pipes for a coming flood of space information. New Earth-observation satellites promise to deliver petabytes of data every year. Missions such as the Mars Reconnaissance Orbiter (MRO) already have constraints on the volume of data they can send back because of fluctuations in download rates tied to a spacecraft's varying distance from Earth.

"Right now, we're really far from Earth, so we can't fit as many images in our downlink," says Ingrid Daubar, who works on the MRO's HiRISE camera at the University of Arizona in Tucson. Laser data highways could ultimately allow space agencies to kit their spacecraft with more sophisticated equipment, says John Keller, deputy project scientist for NASA's Lunar Reconnaissance Orbiter (LRO). That is not yet possible, he says. "We're limited by the rate at which we can download the data."

Today's spacecraft send and receive messages using radio waves. The frequencies used are hundreds of times higher than those put out by music stations on Earth and can cram in

more information, allowing orbital broadcasts to transmit hundreds of megabits of information per second. Lasers, which operate at higher frequencies still, can reach gigabits per second (see 'Tuned in'). And unlike the radio portion of the electromagnetic spectrum, which is crowded and carefully apportioned, optical wavelengths are underused and unregulated.

Efforts to develop laser communication systems struggled for much of the twentieth century: weak lasers and problematic detectors derailed project after project. But recent advances in optics have begun to change the situation. "The technology has matured," says Frank Heine, chief scientist at Tesat-Spacecom, a company based in Backnang, Germany.

In the 1980s, Europe took advantage of improved lasers and optical detectors to begin work on its first laser communication system, the Semiconductor Laser Intersatellite Link Experiment (SILEX). Equipped with the system, the ESA satellite Artemis received 50 megabits of information per second from a French satellite in 2001 and then exchanged messages with a Japanese satellite in 2005. The project taught engineers how to stabilize and point a laser in space. But it was abandoned after its intended application — a constellation of satellites to provide Internet services — was dropped in favour of the network of fibre-optic cables now criss-crossing the globe.

Since then, Heine's team at Tesat-Spacecom has created a laser terminal for satellite-to-satellite communication, at a cost to the German Aerospace Center of €95 million (US\$124 million). The laser, amplified by modern fibre-optic technology, achieves a power of watts — compared with the tens of milliwatts reached by SILEX. In 2008, terminals mounted on two satellites transferred information at gigabits per second over a few thousand kilometres.

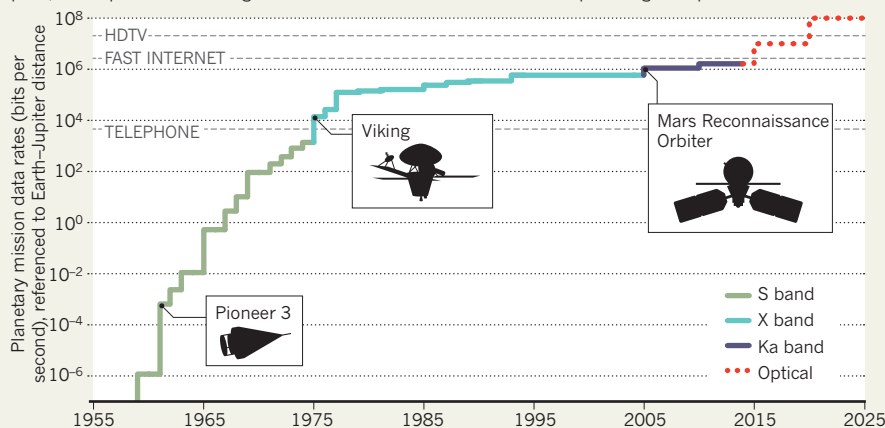
ESA's Alphasat will extend the range of this laser terminal to tens of thousands of kilometres once it is positioned high in geostationary orbit. Future satellites that sport laser terminals in lower orbits will be able to beam as much as 1.8 gigabits per second of information up to Alphasat, which will then relay the data to the ground using radio waves. Alphasat's geostationary orbit means that it can provide

"Laser communication becomes more advantageous the farther out you go."

SOURCE: NASA/JPL-CALTECH

TUNED IN

Interplanetary data transmission rates have shot up 10 orders of magnitude in the past 50 years, thanks in part to higher frequency bands of radio waves. Optical transmissions with lasers promise to extend that pace, to the point at which high-definition television broadcasts from Jupiter might be possible.



a constant flow of data to its ground station — unlike low-Earth-orbit satellites, which can communicate with the ground for only an hour or two each day as they race by overhead. “Other satellites will be able to buy time on our laser terminal,” says Philippe Sivac, Alphasat’s acting project manager.

One client will be another ESA mission due to launch this year: Sentinel-1, the first of several spacecraft to be sent up for Europe’s new global environmental-monitoring programme Copernicus. It will beam weather data to Alphasat until the end of 2014. At that point, Europe plans to start deploying a network of dedicated laser-relay satellites that will ultimately handle 6 terabytes of images, surface-temperature measurements and other data collected every day by a fleet of Sentinel spacecraft.

But Europe’s space lasers have a significant drawback. Although they can shuttle information between spacecraft, they have trouble talking to the ground — a task that must still be performed by radio waves. This is because these lasers encode information by slightly varying the frequency of light in a way analogous to modulating an FM radio station. A beam modulated in this way is protected from solar interference but is vulnerable to atmospheric turbulence.

The laser on NASA’s upcoming LADEE mission will communicate directly with Earth using a different approach that is less susceptible to atmospheric interference. It encodes information AM-style by tweaking the amplitudes rather than the frequency of a light wave’s peaks.

NASA hopes that the LADEE demonstration will extend laser communications beyond Earth’s immediate vicinity, to the Moon and other planets. Deep-space missions currently rely on radio transmissions. But radio waves spread out when they travel long distances, weakening the signal and reducing the data-transfer rate.

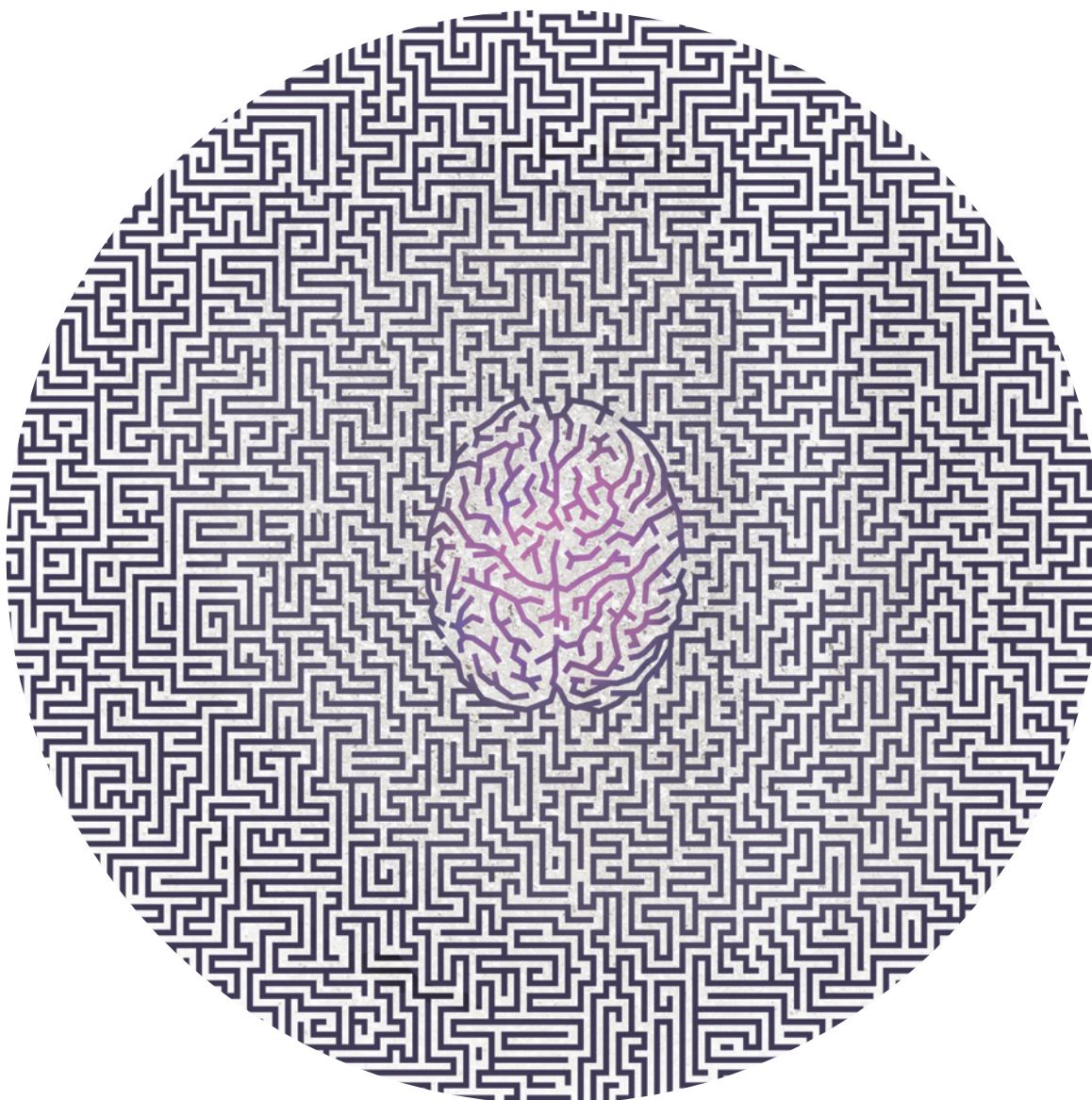
Laser beams, by contrast, keep their focus, allowing them to shuttle the already greater quantities of information they encode over longer distances without using the extra power needed by radio transmitters. “Laser communication becomes more advantageous the farther out you go,” says Donald Cornwell, mission manager for the Lunar Laser Communication Demonstration project on LADEE at NASA’s Goddard Space Flight Center in Greenbelt, Maryland.

In 1992, the Galileo probe, on its way to Jupiter, spotted laser pulses sent more than 6 million kilometres from Earth. A laser on Earth pinged the Mars Global Surveyor in 2005. Another struck the MESSENGER mission en route to Mercury, which responded with its own laser pulses. In January this year, the Lunar Reconnaissance Orbiter received the first primitive message sent by laser to the Moon — an image of the Mona Lisa that travelled pixel by pixel in a sort of Morse code.

LADEE carries NASA’s first dedicated laser communications system. With a bandwidth of 622 megabits per second, more than six times what is possible with radio from the distance of the Moon, the system can broadcast high-definition television-quality video. But even though its AM optical system is good at penetrating Earth’s turbulent atmosphere, it will still need a backup radio link for cloudy days when the laser is blocked. To minimize this problem, LADEE’s primary ground station is in a largely cloudless desert in New Mexico, with alternative sites in two other sunny spots: California and the Canary Islands. ■ [SEE EDITORIAL P.254](#)

CORRECTION

The News story ‘Teething troubles at huge telescope’ (*Nature* **499**, 133–134; 2013) mistakenly gave Natural Resources Canada as the source for the graph instead of National Research Council Canada.



Solving the brain

The United States and Europe are both planning billion-dollar investments to understand how the brain works. But the technological challenges are vast.

BY ALISON ABBOTT

When neurobiologist Bill Newsome got a phone call from Francis Collins in March, his first reaction was one of dismay. The director of the US National Institutes of Health had contacted him out of the blue to ask if he would co-chair a rapid planning effort for a ten-year assault on how the brain works. To Newsome, that sounded like the sort of thankless, amorphous and onerous task that would ruin a good summer. But after turning it over in his mind for 24 hours, his dismay gave way to enthusiasm. "The timing is right," says

Newsome, who is based at Stanford University School of Medicine in California. He accepted the task. “The brain is the intellectual excitement for the twenty-first century.”

It helped that the request for the brain onslaught was actually coming from Collins’s ultimate boss — US President Barack Obama. Just two weeks after that call, on 2 April, Obama announced a US\$100-million initial investment to launch the BRAIN Initiative, a research effort expected to eventually cost perhaps ten times that amount. The European Commission has equal ambitions. On 28 January, it announced that it would launch the flagship Human Brain Project with a 2013 budget of €54 million (US\$69 million), and contribute to its projected billion-euro funding over the next ten years (see *Nature* **482**, 456–458; 2012).

Although the aims of the two projects differ, both are, in effect, bold bids for the neuroscientist’s ultimate challenge: to work out exactly how the billions of neurons and trillions of connections, or synapses, in the human brain organize themselves into working neural circuits that allow us to fall in love, go to war, solve mathematical theorems or write poetry. What’s more, researchers want to understand the ways in which brain circuitry changes — through the constant growth and retreat of synapses — as life rolls by.

Reaching this goal will require innovative new technologies, ranging from nanotechnologies to genetics to optics, that can capture the electrical activity coursing through neurons, prod those neurons to find out what they do, map the underlying anatomical circuits in fine detail and process the exabytes of information all this work will spit out. “Think about it,” says neuroscientist Konrad Kording of Northwestern University in Chicago, Illinois. “The human brain produces in 30 seconds as much data as the Hubble Space Telescope has produced in its lifetime.”

Researchers are already chipping away at the problem: the past few years have seen startling advances in techniques that allow the stimulation of precise neurons deep in the brain using light, for example, and the construction of anatomical maps with unprecedented detail. So far, most neuroscientists are using simpler species such as mice or worms to learn basic principles that evolution may have conserved in humans. Here, *Nature* examines some of the technological advances that will be necessary to drive further, and faster, into the brain.

1 Measuring it

If researchers are to make sense of the frenzy of electrical signals coursing through the brain’s circuits, they will need to record simultaneously from as many neurons as possible.

Today, they typically gauge neuronal activity by inserting metal electrodes into the brain, but this approach comes with enormous challenges. Each electrode needs its own wire to carry out the measured analogue signal — the voltage change — and the signals can easily be lost or distorted as they travel along the wire to instruments that will convert them into the digital signals needed for analysis. Moreover, the wires must be hair-thin to avoid tissue damage. Advances in electrode technologies have seen the number of neurons that researchers can record from double roughly every seven years over the past five decades, such that probes can now reach a couple of hundred neurons simultaneously¹. But the ultimate challenge will require them to reach many more cells and to record higher-quality signals.

That is becoming possible with a new generation of neuroprobes made from silicon, which allows extreme miniaturization. Instruments such as analogue-to-digital converters can be carved out of the same tiny piece of silicon as the electrodes, so the vulnerable analogue signal does not have to travel. A prototype ‘neuroprobe’ of this type was unveiled in February at the International Solid-State Circuits Conference in San Francisco, California, by imec, a nanoelectronics research organization based in Leuven, Belgium. One-centimetre long and as thin as a dollar bill, the probe packs in 52 thin wires and switches that neuroscientists can flip seamlessly between 456 silicon electrodes.

When inserted into a mouse brain, for example, the electrodes dotted

across the imec probe can span — and record from — all layers of the animal’s brain simultaneously, from the cortex to the thalamus in the brainstem. This could help neuroscientists to unpick the circuitry that connects them. “This prototype can be scaled up,” says Peter Peumans, director of bio- and nanoelectronics at imec. Within three years, he says, the neuroprobes will have up to 2,000 electrodes and more than 200 wires.

But rather than just passively measuring electrical activity in neural circuits, researchers also want to test what those circuits do by activating them and observing

the changes that occur in electrical activity and animal behaviour. Each imec probe includes four stimulating electrodes, and in future models this will be increased to 20 or more. But as recording and stimulating electrodes can interfere with one another, researchers are also trying to stimulate neurons with light instead of electricity. These ‘optogenetic’ techniques generally involve inserting light-sensitive ion-channel proteins called opsins into neurons, so that a laser light shone into

“The human brain produces in 30 seconds as much data as the Hubble Space Telescope has produced in its lifetime.”

the skull through an optic fibre opens the channels and activates the neurons. One research group recently used optogenetics in mice, for example, to produce repetitive behaviours that are considered to be a model for obsessive-compulsive disorder².

The next generation of optogenetic neuroprobes will include systems that are able to deliver light directly into the brain exactly where it is needed, without the need for cumbersome optical fibres. In April, for example, Michael Bruchas at Washington University in St Louis, Missouri, and his team described their wireless prototype: an optogenetic chip with light-emitting diodes that can be activated by a radio signal to trigger the opsins³. When the team implanted a chip into mice that could stimulate the brain’s reward centre, the animals quickly learned to switch it on themselves by poking their noses into a hole — showing that the chip worked and could change behaviour.

The search is on for other natural or genetically engineered opsins that respond to different wavelengths of light and might allow researchers to activate and test various elements of a circuit. Eventually, neuroprobes may not only routinely record from and stimulate hundreds or thousands of neurons in mice and non-human primates, but also include sensors to identify neurotransmitters and measure physiological parameters such as temperature, which can affect neuronal activity.

And the future could bring much more radical methods. Some scientists have proposed the idea of nanometre-scale light-sensitive devices that could embed themselves in the membranes of neurons, power themselves from cellular energy and wirelessly convey the activity of millions of neurons simultaneously⁴.

Another idea is to do away with measuring devices and instead capture the post-mortem trace left by an action potential as it passes through the brain. Kording is part of a team trying to do this by exploiting DNA polymerase, the enzyme that cells use to build DNA from its component bases. He and his colleagues have designed a synthetic DNA polymerase that, when surrounded by high levels of calcium, inserts the wrong base into the artificial DNA strand it constructs⁵. If this polymerase could be added to neurons, then an action potential, which causes intracellular calcium levels to spike, would trigger errors in the DNA strand, and the time that this occurred could be determined retrospectively from the length and sequence of the DNA. That’s the theory, anyway, says Kording. “But we are only getting started.”

2 Mapping it

However researchers go about collecting information about neuronal activity and circuitry, it will be essential to map this onto a reliable and highly detailed anatomical atlas of the brain. It is like trying to understand traffic flow in a city: the better the map (the anatomy), the better the predictions of how it will change during rush hour (the active circuits).

For more than a century, the method used to map neuroanatomy has been to slice a brain as thinly as possible, stain the slices to render the cells visible and look at them under the light microscope. But, computationally, it is extremely challenging to align large numbers of slices in order to reconstruct the tangled web of neurons densely packed into a human brain.

Even so, Katrin Amunts of the Research Centre Jülich in Germany and her team announced that they had done it last month, when they published a three-dimensional reconstruction of a human brain in unprecedented detail. To build it, they painstakingly sliced the brain of a 65-year-old woman into 7,400 layers 20 micrometres thick, stained them, imaged them with a light microscope and then used 1,000 hours on two supercomputers to piece the terabyte of data together⁶. The atlas reveals in detail folds of the human brain, which tend to get lost in two-dimensional cross-sections. The whole project took a decade, says Amunts, who has already started work on a second human brain to look at variation between individuals — a project she expects to move a lot faster.

Attempting to take another leap farther, Jeff Lichtman at Harvard University in Cambridge, Massachusetts, and Winfried Denk of the Max Planck Institute for Neurobiology in Munich, Germany, are working with the German optics company Carl Zeiss on a new electron microscope that would image even thinner slices — 25 nanometres, or one-thousandth the thickness of an average cell. “Then you get to see every little damn thing in the brain, from every neuron to every subcellular organelle, from every synapse to every spine neck — everything,” says Lichtman.

Using conventional electron microscopes, with their single scanning beam of electrons, researchers have so far been able to reconstruct only a cubic millimetre of brain tissue. It would take many decades to scan a whole mouse brain's worth of ultra-thin slices, says Denk. The new machines, which should be delivered to the two labs next year, will have 61 scanning beams operating in parallel and will shrink this time down to months. Denk estimates that this will allow them to make a computational reconstruction — “a mouse brain in a box”, as he puts it — within five years.

What Lichtman and Denk have not yet resolved is how to reconstruct a full three-dimensional picture of the tissue from these images. In a trial project using a conventional electron microscope, Denk's lab scanned minuscule volumes of mouse retina, one of the simplest parts of the mammalian brain^{7,8}. But computing alone was not able to reconstruct the 300 gigabytes of image data the effort generated, so the lab enrolled 230 people to help to trace, by eye, the neurons as they meander through the slices. “It won't be practical to do that sort of crowd-sourcing on a larger scale,” says Denk. “We'll have to develop algorithms to get machines to do the job as well as the human eye.”

There may be easier ways to carry out brain mapping at lower resolutions. One possibility is a technique called CLARITY, which generated excitement when it was unveiled in April. Karl Deisseroth at Stanford University and his colleagues have developed a way to chemically replace the opaque lipids in the brain with a clear gel, rendering the tissue transparent and allowing the internal arrangements of neurons to be viewed without the need for slicing⁹. Deisseroth has already applied the technique to brain tissue from a boy who had autism spectrum disorder, and found unusual ladder-like arrangements of neurons in his cortex. Other researchers are clamouring to use the method to trace circuitry in normal brains (see *Nature* **497**, 550–552; 2013).

And however efficient the various activity-measuring and

anatomy-mapping techniques turn out to be, many researchers hope that it won't be necessary to view — or record from — every individual neuron to get a working picture of the whole brain. “Patterns will emerge from which it will be possible to extrapolate,” says Newsome.

3 Making sense of it

Perhaps the most daunting part of the brain challenge lies in storing and handling data. One cubic millimetre of brain tissue will generate an estimated 2,000 terabytes of electron-microscopy information using Lichtman and Denk's new microscope, for example. Denk estimates that an entire mouse brain could produce 60 petabytes and a human brain about 200 exabytes. This amount of data will rival the entire digital content of today's world, “including Facebook and all the big data stores”, says Lichtman.

That is just the start. Neuroscientists will eventually want to collect this type of anatomical information for many human brains — each of them unique — and layer onto it information about neuronal activity. They will need to store and organize all these diverse data types so that scientists can interface with them.

Europe's Human Brain Project, which aims to provide a brain simulation that researchers can interact with in real time, adds another level of demand. “One of our challenges is to develop computer languages that allow a supercomputer's capacity to be used efficiently,” says Jesus Labarta Mancho of the Barcelona Supercomputing Center in Spain, which is a partner of the Human Brain Project. Current supercomputers would be overwhelmed by experiments requiring different parts of the brain to be simulated in different fractions of a second. So the idea is to develop ways to allow the supercomputer to compress information about some brain areas, freeing up resources for computation on the ones that are relevant to the problem at hand.

Even assuming that the data can be neatly packaged, theorists will have to work out what questions to ask of it. “It is a chicken and egg situation,” says theoretical neuroscientist Christian Machens at the Champalimaud Centre for the Unknown in Lisbon. “Once we know how the brain works, we'll know how to look at the data.”

Theorists argue about the scale of the task ahead of them; Kording is one of many who think it will be horrendous. “It makes Google's search problems look like child's play,” he says. “There are approximately the same number of neurons as Internet pages, but whereas Internet pages only link to a couple of others in a linear way, each neuron links to thousands of others — and does so in a non-linear way.”

But Partha Mitra, a biomathematician at Cold Spring Harbor Laboratory in New York, thinks that the bigger challenge to knowing the brain will be sociological. “Chasing after the workings of the brain is not like chasing after the Higgs boson, where everyone goes after the same single target,” he says. “It is about the community setting goals in a deliberate manner and working towards them in a disciplined manner.”

Setting those goals is now consuming Newsome's summer, just as he predicted. He is taking part in a series of expert workshops to define the goals of the BRAIN Initiative and shaping a report on it that is due in September. The report won't promise to solve all the challenges of the brain, he says, but it will set a course that, in the long term, just might.

“We'll eventually learn what all the twinkling of the neurons means in terms of our behaviour,” says Newsome, “and that's what really matters.” ■ SEE EDITORIAL P.253

Alison Abbott is Nature's senior European correspondent.

1. Stevenson, I. H. & Kording, K. P. *Nature Neurosci.* **14**, 139–142 (2011).
2. Ahmari, S. E. *et al. Science* **340**, 1234–1239 (2013).
3. Kim, T.-I. *et al. Science* **340**, 211–216 (2013).
4. Alivisatos, A. P. *et al. ACS Nano* **7**, 1850–1866 (2013).
5. Zamft, B. M. *et al. PLoS ONE* **7**, e43876 (2012).
6. Amunts, K. *et al. Science* **340**, 1472–1475 (2013).
7. Briggman, K. L., Helmstaedter, M. & Denk, W. *Nature* **471**, 183–188 (2011).
8. Helmstaedter, M. *et al. Nature* (in the press).
9. Chung, K. & Deisseroth, K. *Nature Meth.* **10**, 508–513 (2013).

COMMENT

DIGITAL LEARNING Massive open online courses can stimulate research **p.277**



EVOLUTION Machiavelli's work *The Prince* speaks to today's behavioural sciences **p.280**

GENOMICS Seven-storey art installation at the Broad Institute **p.283**

REPRODUCIBILITY Two more red flags for suspect work involving big data **p.284**

ILLUSTRATIONS BY CHRIS RYAN/NATURE



Look, then leap

Massive open online courses can make higher education more accessible, immersive and comprehensive — if they are deployed with due caution, says **Michael M. Crow**.

The response by many in academia and the media to massive open online courses (MOOCs) that make content freely accessible to millions of learners is reminiscent of the hysteria with which alien invaders were met in the pulp science-fiction films of the 1950s. Folks fear the unknown.

More reasoned reception was expressed in a highly publicized open letter sent in April from members of the philosophy department at San José State University in California to Michael Sandel, a professor

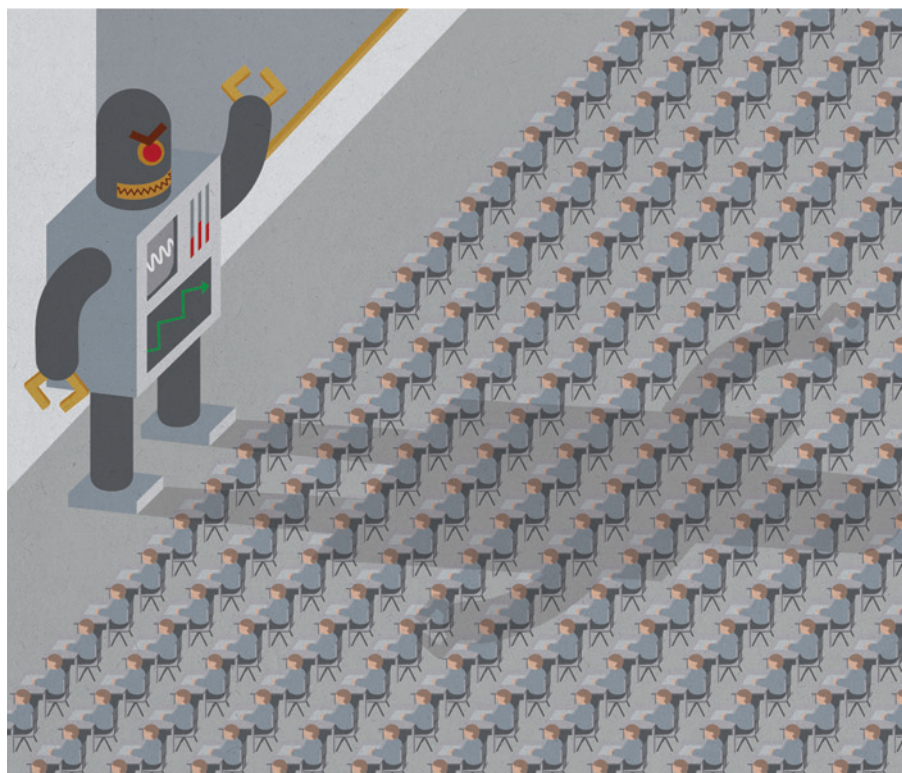
of government at Harvard University in Cambridge, Massachusetts (see go.nature.com/7n9tgi). The letter explained a refusal to pilot Sandel's interactive online course on social justice. The 'blended' course, which combined online learning with classroom teaching, was proffered by edX, the non-profit

provider of MOOCs established by Harvard and the Massachusetts Institute of Technology in Cambridge. The San José philosophers stated that the "move to MOOCs comes at great peril to our university" and threatens "serious compromise of quality of education". In May, 58 members of the Harvard University Faculty of Arts and Sciences expressed similar scepticism over MOOCs in a letter to their dean, saying that they were "deeply concerned" about the "impact online courses will have on the higher education system as a ▶



DIGITAL LEARNING

A *Nature* and *Scientific American* special. nature.com/digitallearning



▶ whole” (see go.nature.com/tiru87).

What these letters identify is the lingering ambivalence in academia towards the transition from teaching and learning, based on the fifteenth-century technology of the printed word, to twenty-first-century interactive technologies that offer the potential for adaptive, personalized learning on an infinite scale. Society has embraced the Internet as axiomatic to individualized learning, personal empowerment, cultural evolution and, more recently, revolution. People born after 1990 have never known life without the Internet; for them, it is as integral to learning as it is to their social lives. All aspects of education — from information acquisition to adaptive learning — have been forever altered by a technology that only 20 years ago was considered by nationally prominent academics to be a passing craze.

The pace of technological change has generated much soul-searching about how Internet-delivered courses will change the art of learning. Some worry that a single professor on a computer will, for instance, teach all of the microeconomics courses in an entire country. Others imagine that a latter-day Charles Darwin will, by his or her genius, transform the teaching of introductory university-level biology. But the art of learning has already changed completely, because for almost a decade students have had instant access to unlimited information from anywhere or anyone in the world. This has altered all assumptions about academic hierarchy, charismatic authority, pedagogical processes and the structure of the learning community.

Those in academia are still processing the implications of the impact of digital learning on the diverse arena of higher education, in which students from every socioeconomic background attend institutions that range from community colleges to the world's leading research universities. As colleges and universities educate leaders and experts in every sphere of human endeavour and conduct a large share of the world's scientific research, we must lead and shape this revolution, not recoil from it, if we are to avoid stifling innovation.

“We must lead and shape this revolution, not recoil from it, if we are to avoid stifling innovation.”

FROM FAD TO REVOLUTION

MOOCs have been on the horizon for two decades. In 1993, when I was a faculty member and vice-provost for research at Columbia University in New York, I hosted a meeting for visitors from the University of Illinois who had developed a tool to link computers on different systems and in different places. One of the visitors was Marc Andreessen, then a student and technical assistant. Andreessen went on to co-develop Mosaic, the first widely popular Web browser, which helped to facilitate the Internet boom of the 1990s. Most of the Columbia faculty members attending the presentation dismissed the Internet as a fad. A few of us suspected that this technological innovation might empower and accelerate the creation and dissemination of knowledge.

During the next few years, I oversaw investment by Columbia University in dozens of faculty groups focused on digital learning. In 2000, we launched Fathom.com, an interactive platform offering free academic content and hundreds of courses provided by a consortium of around 15 of the world's leading universities and cultural institutions — arguably an early precursor to MOOCs. By 2002, around 45,000 students had registered for one of Fathom's online courses.

The platform remained online for roughly three years, but it was slow to generate revenue, heavily constrained by the bandwidth limits of the Internet at the time and by cultural confusion as to its purpose, even among members. After all, wasn't learning supposed to be conducted face to face, in environments rich with discourse? Back then, this slow, choppy and often unreliable thing called the Internet lacked sufficient interactive capacity.

Two decades on from that meeting at Columbia, and more than ten years on from Fathom's launch, it baffles me that a similar scepticism of Internet-enhanced learning still thrives among some faculty members, even as some of the world's leading universities embrace online distance learning as a useful platform and millions of learners worldwide enrol in MOOCs. But an online course developed by faculty members and offered through an online degree programme is not the same thing as a MOOC, and we must observe this distinction if we are to gain understanding about their impact.

For students fortunate enough to enrol in research universities and liberal-arts colleges, undergraduate educations are best advanced in a milieu of discovery and learning immersion. If conceived and executed properly, online courses could provide an important complement. For other students — those living in isolated areas, say, or studying later in life — online courses or fully online degree programmes offer an alternative to the traditional undergraduate experience. Notwithstanding that a course is not a curriculum, that a curriculum is not a degree, and that a degree does not automatically produce an enabled and adaptive learner, our evolving capacity to improve on the textbooks of the past is as good for the faculty members and students of a community college as it is for those in leading universities.

The revolutionary aspect of MOOCs is their potential to reach millions of learners who are not enrolled in colleges and universities. It seems ironic that the very institutions that have become the most exclusive in terms of admissions practices and influential in terms of research have taken the lead in offering MOOCs developed by their professors to millions of students. In fact, this transition from limited access to mass-market is consistent with these universities' values: it

expands their societal impact. I believe that online learning will enable the creation of high-speed and possibly more efficacious multi- and interdisciplinary teaching environments around the world.

CASE BY CASE

As president of Arizona State University (ASU) — a large public research university that has four campuses in the Phoenix metropolitan area — I am optimistic about the potential for online learning to enable broader access to the highest levels of academic excellence. The ASU has embraced online learning and offers fully online undergraduate and graduate degree programmes. Faculty members are in charge of course content and complement online instruction with personal interaction. Interactive, adaptive learning platforms break down content and present it to students at an optimal pace and in a tailored sequence, while giving detailed feedback on performance to both students and professors. Our objective is to create an environment in which a person can learn anything, and thus we must avail ourselves of every tool, mechanism and means. The challenge remains for the ASU, like all universities, to evaluate the relevance of MOOCs to curricula and degree programmes and to determine whether and how to proceed with their implementation. This will inevitably advance on a case-by-case basis.

Regardless of any apprehension about online learning, there is now an opportunity to transcend the limits of the 'sage on the stage'. Online learning — including MOOCs — can enable and accelerate deeper or broader learning in certain subjects and thus



complement face-to-face or team-learning modalities. Where appropriate, MOOCs represent a simple, straightforward way to put quality content in front of lots of students at the same time, enabling scarce resources to be deployed where they are most needed: in the creation of teaching, learning and discovery environments that can educate millions to the levels necessary to ensure successful and continued adaptation in the future.

One cautionary note should be sounded: to the naive, MOOCs seem like the quick fix for cost containment — itself a crucial objective if there is to be any social equity in the higher-education system. A professor from an exclusive university, such as Columbia, could have his or her course put onto the Web as a MOOC, then students at larger public universities could take that course instead of its equivalent on their home campuses, thereby shrinking faculties and costs. But producing a master adaptive learner will always require interaction with the creative and caring teacher-scholars known as the faculty. The integration of interactive online learning tools such as MOOCs can be useful, but these disembodied teaching vehicles are no substitute for teacher interaction.

I can only hope that in our enthusiasm for MOOCs, we look before we leap into a dystopia in which the masses are taught by robots, and the gifted and wealthy by living, breathing professors. We need to end up in a world in which technology once and for all makes higher education more accessible by lowering its cost and enhancing its impact across our great and highly diverse society. ■

Michael M. Crow is president of Arizona State University, Tempe, Arizona 85287, USA. He previously served as executive vice-provost and professor of science policy at Columbia University in New York. e-mail: michael.crow@asu.edu

Online on-ramps

Giving scientists greater access to conceptual and technical tuition through massive open online courses will aid interdisciplinary research, say **Hazel Sive** and **Sanjay Sarma**.

There has been a lot of discussion about the way in which massive open online courses (MOOCs) might change the landscape of education. A transformation is likely, but in a form more nuanced than current conversations imply (see go.nature.com/cynqvp).

Another question has had less attention: how will MOOCs affect research and scholarly enterprise at universities? Will the courses lead to a decline in the amount or quality of study? Or will a huge increase in online learning improve research?

We are both deeply involved in the MOOC culture at the Massachusetts Institute of Technology (MIT) in Cambridge. We are members of the university's Task Force on the Future of MIT Education, which

focuses on the use and effect of MOOCs, and we both run active research groups, so we appreciate the dual tugs of teaching and scholarship. Our opinion is that, at least for residential, research-focused universities, the main impact of MOOCs on research culture will be to aid interdisciplinary work by enabling greater access to conceptual or technical tuition.

Many factors influence scholarship and research in a university. These include funding, space, equipment, research collaborations, the number of faculty members, the

presence of faculty members on campus and the time available for research. Research is often deeply affected by the insight and questioning of smart undergraduate students or postdocs. Fields move forwards through the ability to think in novel, cross-disciplinary ways, and through familiarity with state-of-the-art techniques.

Which parts of that equation will MOOCs affect? A common view is that they will replace professors. Clearly, this would diminish the volume of research and stifle the collegial discourse that fosters collaboration and original thinking. We do not consider this to be a likely outcome — except perhaps in parts of the world where students have limited access to quality education and where a MOOC might stand in for ▶



DIGITAL LEARNING

A *Nature* and *Scientific American* special. nature.com/digitalllearning

expands their societal impact. I believe that online learning will enable the creation of high-speed and possibly more efficacious multi- and interdisciplinary teaching environments around the world.

CASE BY CASE

As president of Arizona State University (ASU) — a large public research university that has four campuses in the Phoenix metropolitan area — I am optimistic about the potential for online learning to enable broader access to the highest levels of academic excellence. The ASU has embraced online learning and offers fully online undergraduate and graduate degree programmes. Faculty members are in charge of course content and complement online instruction with personal interaction. Interactive, adaptive learning platforms break down content and present it to students at an optimal pace and in a tailored sequence, while giving detailed feedback on performance to both students and professors. Our objective is to create an environment in which a person can learn anything, and thus we must avail ourselves of every tool, mechanism and means. The challenge remains for the ASU, like all universities, to evaluate the relevance of MOOCs to curricula and degree programmes and to determine whether and how to proceed with their implementation. This will inevitably advance on a case-by-case basis.

Regardless of any apprehension about online learning, there is now an opportunity to transcend the limits of the 'sage on the stage'. Online learning — including MOOCs — can enable and accelerate deeper or broader learning in certain subjects and thus



complement face-to-face or team-learning modalities. Where appropriate, MOOCs represent a simple, straightforward way to put quality content in front of lots of students at the same time, enabling scarce resources to be deployed where they are most needed: in the creation of teaching, learning and discovery environments that can educate millions to the levels necessary to ensure successful and continued adaptation in the future.

One cautionary note should be sounded: to the naive, MOOCs seem like the quick fix for cost containment — itself a crucial objective if there is to be any social equity in the higher-education system. A professor from an exclusive university, such as Columbia, could have his or her course put onto the Web as a MOOC, then students at larger public universities could take that course instead of its equivalent on their home campuses, thereby shrinking faculties and costs. But producing a master adaptive learner will always require interaction with the creative and caring teacher-scholars known as the faculty. The integration of interactive online learning tools such as MOOCs can be useful, but these disembodied teaching vehicles are no substitute for teacher interaction.

I can only hope that in our enthusiasm for MOOCs, we look before we leap into a dystopia in which the masses are taught by robots, and the gifted and wealthy by living, breathing professors. We need to end up in a world in which technology once and for all makes higher education more accessible by lowering its cost and enhancing its impact across our great and highly diverse society. ■

Michael M. Crow is president of Arizona State University, Tempe, Arizona 85287, USA. He previously served as executive vice-provost and professor of science policy at Columbia University in New York. e-mail: michael.crow@asu.edu

Online on-ramps

Giving scientists greater access to conceptual and technical tuition through massive open online courses will aid interdisciplinary research, say **Hazel Sive** and **Sanjay Sarma**.

There has been a lot of discussion about the way in which massive open online courses (MOOCs) might change the landscape of education. A transformation is likely, but in a form more nuanced than current conversations imply (see go.nature.com/cynqvp).

Another question has had less attention: how will MOOCs affect research and scholarly enterprise at universities? Will the courses lead to a decline in the amount or quality of study? Or will a huge increase in online learning improve research?

We are both deeply involved in the MOOC culture at the Massachusetts Institute of Technology (MIT) in Cambridge. We are members of the university's Task Force on the Future of MIT Education, which

focuses on the use and effect of MOOCs, and we both run active research groups, so we appreciate the dual tugs of teaching and scholarship. Our opinion is that, at least for residential, research-focused universities, the main impact of MOOCs on research culture will be to aid interdisciplinary work by enabling greater access to conceptual or technical tuition.

Many factors influence scholarship and research in a university. These include funding, space, equipment, research collaborations, the number of faculty members, the

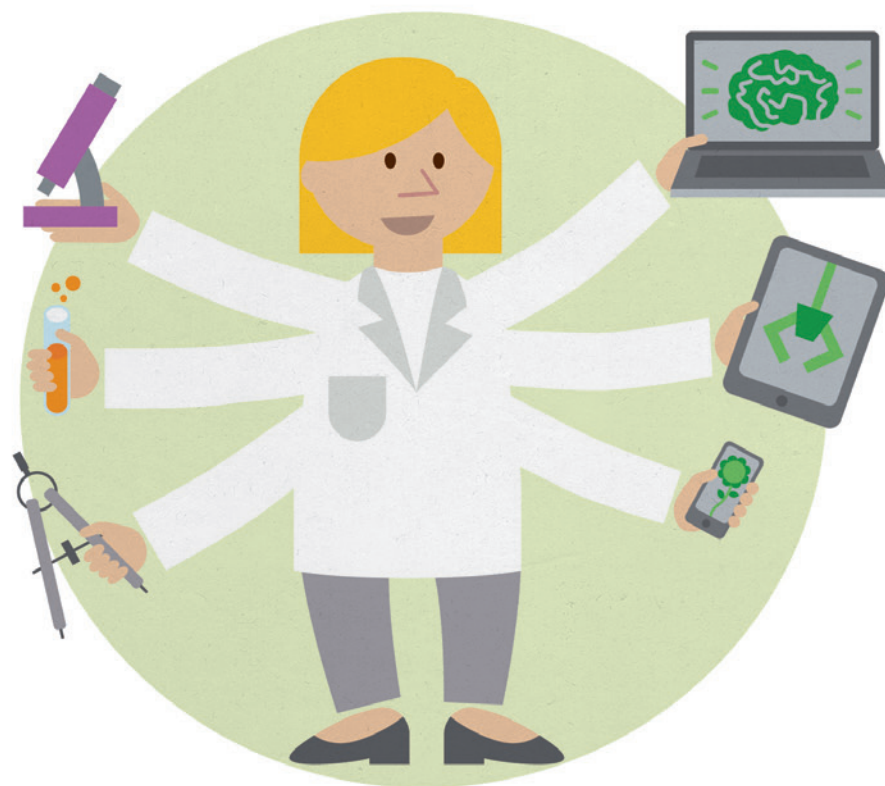
presence of faculty members on campus and the time available for research. Research is often deeply affected by the insight and questioning of smart undergraduate students or postdocs. Fields move forwards through the ability to think in novel, cross-disciplinary ways, and through familiarity with state-of-the-art techniques.

Which parts of that equation will MOOCs affect? A common view is that they will replace professors. Clearly, this would diminish the volume of research and stifle the collegial discourse that fosters collaboration and original thinking. We do not consider this to be a likely outcome — except perhaps in parts of the world where students have limited access to quality education and where a MOOC might stand in for ▶



DIGITAL LEARNING

A *Nature* and *Scientific American* special. nature.com/digitallearning



► face-to-face teaching, possibly leading to a cut in faculty numbers. Data on this point are not yet available owing to the newness of these teaching tools.

In research-intensive universities, it is our experience that the time in a faculty member's schedule that is freed up through the use of online material is invested back into 'flipped' classrooms — students have already studied lecture material online, meaning that class time is used for discussion, hands-on activities, problem solving and research. Similarly, cutting-edge graduate classes thrive on faculty–student interaction. Any reduction in lecture time is reinvested in the classroom to make the experience more comprehensive and fruitful for graduate students.

When online learning tools were used to teach electrical engineering to MIT undergraduates, faculty time that would have been devoted to lecturing went back into the course as alternative ways of teaching the material. These included practical demonstrations, hands-on experiments and interactive laboratory sessions featuring design, creation, debugging, small group discussions and one-to-one student–instructor feedback.

The increase in problem solving and discussion could expose more undergraduates to research questions and strategy earlier and more extensively than at present. This could encourage undergraduates to pursue research during their degree period or beyond, thereby enriching research culture.

For faculty members devising MOOCs and developing additional learning tools, there might be a short-term dip in research

productivity. Such a dip is experienced by any devoted teacher who devises a new course, but the high-tech nature of MOOCs can involve a bigger time investment than normal. Our colleagues at MIT report that developing a new MOOC takes between half a semester and a full one, which lasts about 14 weeks.

The most important impact on research will probably be interdisciplinary online education. Say that an MIT student needs to learn current techniques in machine learning to apply to her main area of study, mechanical engineering. An appropriate class is not offered at MIT until the next spring, but the University of California, Berkeley, offers an online course that she can take this summer. She takes the course and an adviser at MIT vets her progress. By the autumn, the student is ready to apply her newly acquired skills to her thesis research.

A related and crucial benefit of MOOCs to research stems from modules that will help investigators to straddle fields by aiding fluid learning. Modules can range from, for example, a 10-minute video segment to the content of multiple full-length lectures. Thus, for example, a postdoctoral neuroscientist might need to learn how to measure changes in cell shape. A full course is not necessary. Multiple modules explaining relevant concepts are available online from

several universities. The postdoc uses these to increase his competency faster, more cheaply and more comprehensively than he would by taking a course, visiting a relevant laboratory or attending a meeting. This sort of learning has the potential to improve research output.

The serious decline of research funding in the United States and in other parts of the world is reducing the ability of researchers to travel to acquire training in person. Although we in no way suggest that virtual education can entirely plug this gap, it can help scientists to gain some of the conceptual and technical expertise that is necessary to advance research.

Although we feel that online education is unlikely to dent faculty engagement in research, it will be important to monitor carefully the amount of time that faculty members spend on developing MOOCs. We suggest that this is done at the departmental level to ensure that a push for online educational development is balanced by responsiveness to any research challenges that result. ■

Hazel Sive is professor of biology and associate dean of science at the Massachusetts Institute of Technology (MIT) in Cambridge, Massachusetts, USA. She is also at the Whitehead Institute for Biomedical Research, Cambridge, Massachusetts. **Sanjay Sarma** is professor of mechanical engineering and director of digital learning at MIT.
e-mails: sive@wi.mit.edu; sesarma@mit.edu

ILLUSTRATION BY CHRIS RYAN/NATURE

IN RETROSPECT

The Prince

John Whitfield finds resonance with today's behavioural sciences in Niccolò Machiavelli's great Renaissance political treatise, begun five centuries ago this month.



Machiavelli knew that an individual's fate is inseparable from that of his or her group's.

As a high-ranking minister and diplomat in the Florentine republic, Niccolò Machiavelli might have been surprised by his enduring reputation as the prophet of a philosophy in which winning is all and scruples are for suckers. His classic work *The Prince*, written 500 years ago, is a foundational text of political science with continuing relevance for one of today's hot research topics. Since its publication five years after the author's death, it has also been infamous for severing the link between politics and morality.

The Prince can be read both as a guide to how men acquire and maintain power and a book about how societies work — how states can create the public spirit they need to survive. Today, this is the point where disciplines such as game theory, evolutionary psychology and behavioural economics meet, and where researchers still argue about human nature and how their academic insights can translate into public policy.

Machiavelli based his case on observations made in the courts of Europe during 14 years in government, as well as readings in classical history. *The Prince* was a revolutionary twist on the medieval and Renaissance genre known as 'mirrors of princes', offering advice to rulers: it contains few references to Christian virtue, instead emphasizing evidence-based policy. Chapter by chapter, the book guides the neophyte dictator through problems such as the best way to treat conquered states, how to plan for a siege and the right image to project to the masses.

Machiavelli is particularly interested in upstart rulers who have risen through daring and force. If *The Prince* has a hero, it is Cesare Borgia, who in around 1500 cut a swathe across Italy using murder and trickery. Borgia also temporarily unified a good part of the peninsula under a native-born ruler — and so, in Machiavelli's eyes, achieved more good than many more conventionally virtuous leaders.

This tough-mindedness reflects the turbulence of Machiavelli's time in government, which began in 1498. He saw Florence and other Italian states divided by conflict and invaded by foreign powers such as Spain and France. His political career ended in 1512, when the Medici family and the 'warrior pope' Julius II retook Florence and abolished the republic. In early 1513, Machiavelli was tortured on suspicion of conspiring against the Medici and then released without charge. In July, he began to write what would be *The Prince*. Some six months later, he presented it to the Medici, as one of history's more remarkable but less successful speculative job applications.

A lot of what we have learned since Machiavelli

The Prince
NICCOLÒ
MACHIABELLI
1532.

THE BRIDGEMAN ART LIBRARY



Punishment is costly and invites retaliation, so why do we do it?

about our species' social behaviour can be summed up by saying that humans' exceptional ability to cooperate with members of their own group gives them an exceptional ability to compete with other groups. This is called parochial altruism. Humans are more helpful to each other than almost any other species, and their societies are more egalitarian. But beyond the edge of a group, cooperation has a tendency to become well-organized brutality.

The Prince is a treatise on parochial altruism. Machiavelli knew that an individual's fate is inseparable from that of his or her group, and had seen that weak states fell prey to those that were stronger and better organized. You don't have to squint too hard to see the book not as a self-help guide for despots, but as an investigation of how societies can build and maintain the public good in an unconducive political environment.

For example, Machiavelli's prescriptions for resisting invasion and suppressing plots focus on building collective strength. Even for an autocrat, he argues, success comes through good government. The "friendship of the people", he says, is a ruler's best insurance against conspiracies, and a citizen militia is a more effective fighting force than foreign mercenaries. He gives more, and harsher, advice on dealing with outsiders — "a Prince", he writes, "must have no other object or thought ... except war" — than he does on domestic policy, where he recommends rewarding talent and enterprise, not overtaxing one's subjects, and staying away from their property and women.

Beyond the city walls, the gloves are off. For example, he says: "You can be more liberal with what does not belong to you or your subjects." But Machiavelli understood

that boundaries can be manipulated. He advises that the best way to hold a conquered territory is to go and live there, so that its group becomes yours. He yearned for Italy's unification more than three centuries before it happened.

Even when Machiavelli famously advises that fear, and not love, is the safest guarantee of loyalty, he highlights the wider benefits of a fearsome reputation. Cruelty should be "as far as possible turned to the good of one's subjects". He writes that "by making an example or two", a ruler "will prove more compassionate than those who, being too compassionate, allow disorders which lead

"The Prince can be read both as a guide to how men acquire and maintain power and a book about how societies work."

to murder and rapine. These nearly always harm the whole community, whereas executions ordered by a prince only affect individuals."

This is similar to current debates over why individuals and societies punish wrongdoers. It is unclear why people are so keen to mete out punishment, which is costly, uses up time and energy, and invites retaliation. And if a punished person reforms, the benefit is spread across the whole community. This creates an incentive to let other people do the punishing.

Researchers disagree over the selective forces that favour punishment. Those with a background in evolutionary biology, such as Manfred Milinski, tend to view it as a tool that primarily serves individuals, by displaying strength, winning friends and deterring challengers. Others, generally economists,

including Ernst Fehr, emphasize the benefits to groups of suppressing antisocial behaviour. In this view, known as strong reciprocity, punishment arose and spread because groups that did it outcompeted groups that did not.

Machiavelli would probably have taken the former, neo-Darwinian, view that to survive, every altruistic act ought to yield an individual benefit. Despite his republican sympathies, he wrote that men were "ungrateful, fickle, liars and deceivers", who "shun danger and are greedy for profit".

But believers in strong reciprocity point instead to the wealth of evidence that instincts for fairness, sharing and justice are deeply rooted in human behaviour. In a 2010 lecture series (due to appear as a book) at Yale University titled 'Machiavelli's mistake', one of them — economist Samuel Bowles of the Santa Fe Institute in New Mexico — argues that when governments, assuming people to be self-interested, focus on material incentives as a means of encouraging good behaviour, they risk undermining the prosocial behaviours that underpin society.

Given the breadth of Machiavelli's writing, and his diligence as a public servant, it is unjust that his name epitomizes manipulative selfishness. Even so, how we run our societies still depends on understanding to what extent people are naturally Machiavellian and — more importantly — which environments make them more so. ■

John Whitfield is the author of *People Will Talk: The Surprising Science of Reputation*, and comment editor of *Research Fortnight* and *Research Europe*.
e-mail: j.a.whitfield@gmail.com



NEUROPHILOSOPHY

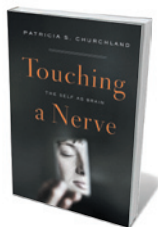
My brain and I

Chris Frith reflects on a book that probes the knotty nexus between brain and mind.

Patricia Churchland is the doyenne of neurophilosophers. She believes, as I do, that to understand the mind, one must understand the brain, using evidence from neuroscience to refine concepts such as free will. Many philosophers and others are unhappy with this proposal. The problem, Churchland writes, is that deep down we are all dualists. Our conscious selves inhabit the world of ideas; our brains, the world of objects.

So deep is this split that we find it hard to accept an intimate relationship between the mind and brain. In *Touching a Nerve*, Churchland hopes to help us overcome this aversion and accept the “neural reality of our mental lives”. To encourage the general reader, she emphasizes her background as an unsophisticated country girl whose common sense stems from growing up on a farm in an isolated valley in British Columbia, Canada.

She begins by showing us how common sense and neuroscience reveal that there is no need for a soul. We are beginning to have an inkling of the underlying mechanisms that enable thinking, feeling and deciding, such as the precise way in which the anaesthetic procaine removes the sensation of pain. Common sense and neuroscience also tell us that there is no life after death. The light at the



Touching a Nerve: The Self as Brain

PATRICIA CHURCHLAND
W. W. Norton: 2013.

end of the tunnel associated with near-death experiences is the effect of oxygen starvation on the brain's visual system. Churchland goes on to discuss morality, aggression, free will and consciousness. But if you were expecting thorough-going interpretations of these concepts in neuroscientific terms, you will be disappointed. She promotes the ‘ordinary’ meaning of free will — “intending your action, knowing what you are doing, and being of a sound mind”. She does not consider the disturbing results of neuroscience research, which suggest that awareness of action — intending and knowing — occur after the action has been selected. We are also told that moral values such as honesty, loyalty and courage depend on learning local conventions and hearing the “stories [that] give you a sense of the right way to act”.

I have no quarrel with the idea that upbringing and culture have important roles in determining behaviour, but this does not seem compatible with Churchland's view that

“our brains determine everything about who we are and how we experience the world”. She also misses the opportunity to present studies that explore the links between brain and culture. There are special processes in the human brain, such as the ability to imitate others with high fidelity, that enable the cumulative development of culture. At the same time, culture moulds the brain and may even drive genetic evolution (see S. E. Fisher & M. Ridley *Science* **340**, 929–930; 2013). Each human brain is part of a dynamic, interacting system of other brains embedded in culture.

What neuroscience there is in *Touching a Nerve* is accurate and commendably up to date. There are useful notes associated with each chapter, including primary sources. Yet I became increasingly irritated by the mixture of science and homespun wisdom. Stories about badly behaved schoolgirls and White Leghorn hens did not help my understanding of the basis of aggression and sex. And the referencing is patchy: why does the statement “not every disappointment can be remedied” deserve a reference, whereas the neural basis of Charles Bonnet syndrome and the claim that patients with schizophrenia can tickle themselves do not? As for common sense, I agree with developmental biologist Lewis Wolpert that the important findings of science typically go against it. It is the data supporting the common-sense interpretation that need to be most carefully checked.

Nevertheless, it may well be true that dualism is deeply ingrained in our nature. A recent brain-imaging study revealed that we have two circumscribed brain circuits: one enables us to think about mental causation, such as how unfairness makes us angry; the other enables us to think about physical causality, such as how heat activates pain receptors. These circuits are mutually antagonistic, so we cannot do both at once (see A. I. Jack *et al.* *NeuroImage* **66**, 385–401; 2013). But if mind–brain dualism is so deeply ingrained, why are the shops full of books such as *Touching a Nerve*, which show that it is the brain that makes decisions, determines moral values and explains political attitudes?

I can only assume that these are the modern equivalent of Gothic horror stories. We love to be frightened by the thought that we are nothing more than the 1.5 kilograms of sentient meat that is our brain, but we don't really believe it. I don't think Churchland really believes it either. ■

Chris Frith is emeritus professor of neuropsychology at the Wellcome Trust Centre for Neuroimaging at University College London, and a fellow of All Souls College, Oxford. His books include *Making Up the Mind: How the Brain Creates Our Mental World*.
e-mail: c.frith@ucl.ac.uk



Fifth-floor view of artist Daniel Kohn's *Instance of a Visual Dataset*, displayed across seven storeys at the Broad Institute in Cambridge, Massachusetts.

GENOMICS

Art of the hypothetical

Elie Dolgin gets the inside story on the data displays rendered as abstract visualizations at a renowned genomics institute.

Exit the main lift on any floor at the Broad Institute's headquarters in Cambridge, Massachusetts, and the first things you will see on the far landing wall are dozens of coloured aluminium squares, each about the size of a pizza box. The tiles in these big, brilliantly hued metallic mosaics at the genomics and biomedical research centre look jumbled, almost random. Gaps litter the scene. Many tiles sit seemingly out of place; others are white. The installation is spread across the building's seven storeys, with themes explored on one level often continuing to the next.

The institute is the setting for an artistic experiment. Each tile shows part of an abstract digitized watercolour painting by artist Daniel Kohn, inspired by fundamental ideas in medical research. So a lattice of red squiggles greets you on the third floor, where most of the chemical biologists work: the interlocking structures are suggestive of the relationships between atoms, covalent bonds and electron orbitals. The seventh floor, home to a hotchpotch of geneticists, cancer biologists, metabolome researchers and more, hosts mostly straight green lines on a white background. The linearity is reminiscent of the way in which scientists interpret the genetic code, with tiles inverted or translocated to evoke genomic drivers of disease.

This installation, entitled *Instance of a*

Visual Dataset, is the culmination of a decade's relationship between the Brooklyn-based artist and the lab. It started in 2003, when Broad Institute co-founder Todd Golub, then an oncologist at the Dana–Farber Cancer Institute in Boston, Massachusetts, admired some of Kohn's work at a hotel. The two struck up a partnership and Kohn served as the institute's first official artist-in-residence from 2006 to 2008, producing smaller works.

Instance of a Visual Dataset was a subsequent commission. To develop his ideas, Kohn met Broad Institute scientists to discuss active research topics, from DNA sequencing to small-molecule drug discovery. He then painted free-form interpretations of what he had learned onto tile grids — usually arranged three by three, but sometimes in other configurations. Kohn digitized these to make more than 3,000 panels, of which about 1,400 are displayed at the institute at any one time. The name of the work describes the 'instantiation' of Kohn's larger data set — its representation at any one moment in the material world.

Since October 2012, Kohn's images have migrated across the installation, mutating on different planes. Kohn has introduced new panels, taken others away and shifted what was already there to find the arrangements

he wanted to leave behind. "It's not just about projecting an image," he says. "It's about listening to where the image takes you."

At the end of June, Kohn adapted the project for the final time. The panels are all now locked in place, where they will remain. A formal opening of the installation is scheduled for late September. Meanwhile, Kohn is planning his next academic partnership, with the Albert Einstein College of Medicine's Center for Epigenomics, and is in the process of launching an organization dedicated to promoting collaborative explorations of nature called the Art Science Observatory.

Kohn "expanded my horizons with respect to things I work with every day," says Damian Young, a chemist at the Broad Institute. "It will be interesting to see," he adds, "how those paintings grow in their meaning and what collaborations they spawn by allowing scientists to conceptualize different things."

With so many ideas swirling around together throughout the building, it can be difficult to find the central theme of Kohn's work. But therein lies his vision. "The piece," Kohn says, "is really about how divergent data can be brought together into a coherent framework — which is ultimately what researchers are trying to do in science, too." ■

Elie Dolgin is senior news editor of *Nature Medicine* in Cambridge, Massachusetts.

Correspondence

Academy reform needs a reality check

Government reform of the Russian Academy of Sciences is timely (*Nature* **499**, 5–6; 2013). But it must also be realistic. Simply imposing foreign models could be costly, inefficient and have disastrous consequences for the country's scientific community.

The Russian government contends that the academy is inefficient, so we suggest an international audit of its output. The results could then be compared with those of recent, costly government initiatives to commercialize science, such as the nanotechnology firm Rusnano and the Skolkovo Innovation Center near Moscow.

The government has a poor record of agency reforms: at the defence ministry, for example, changes triggered corruption, sale of the property to private investors and degradation of the army. We propose a moratorium on changes to the academy's buildings to prevent its arbitrary closure.

The proposal to merge the academy with other minor academies in Russia, which differ in their aims and scientific potential, could result in a lower scientific status for members and an over-representation of the biomedical sciences. If the merger goes ahead, members would still need to be selected on the basis of scientific merit alone.

The Russian Foundation for Basic Research could be given a major role in funding the merged academy's research, which would be better than creating a relatively unstructured government agency.

In our view, publications and other impact-based criteria used to evaluate the productivity of individual scientists and research teams should involve the Russian and the international scientific community.

A demand for rapid innovation must be underpinned by fundamental science. Lessons can be learned from 1948, when

the government followed up agronomist Trofim Lysenko's empty promises to improve agriculture and effectively destroyed Soviet genetics, one of the world's best genetics schools.

Sergey V. Razin *Institute of Gene Biology, Russian Academy of Sciences; and Lomonosov Moscow State University, Moscow, Russia.*

Yegor S. Vassetzky *Institut Gustave Roussy (CNRS UMR8126), Villejuif, France.* vassetzky@igr.fr

Two more red flags for suspect work

I suggest two more 'red flags' in addition to the six that C. Glenn Begley identifies for evaluating preclinical studies (*Nature* **497**, 433–434; 2013). These extend Begley's question regarding the suitability of statistical tests, and apply particularly to computational analyses of large amounts of data — such as those generated by high-throughput proteomics experiments.

The first new flag concerns the application of a multiple-hypothesis correction. The large number of statistical comparisons made in high-throughput data analyses will inflate estimates of significance by increasing the probability that an individual result at a particular significance could occur by chance.

Essentially, this extends the question, "how likely is it that the difference I observe in one measurement is a chance finding?" to the population-level question "how likely is it that I would find this difference by chance if I were to look at a whole bunch of measurements?" An example would be to assume that the chances of finding a left-handed player on a basketball team would simply be around 10% — the chance that an individual is left-handed. The real probability would be higher because many players are being tested.

The second flag questions

whether an appropriate background distribution was used. It is vital to choose a set of variables appropriate to the question that the experimental results are being tested against for significance. An inappropriate choice of background can artificially induce significance in the results or mask real results. An example would be to sample a women's basketball team to determine whether there is a significant height difference between men and women.

Jason E. McDermott *Pacific Northwest National Laboratory, Washington, USA.* jason.mcdermott@pnnl.gov

Promotional tactics corrupt research

With more than US\$1 trillion spent globally on research and development in 2007 (see go.nature.com/5wdd9p), sheer scale seems to be corrupting the scientific enterprise as individuals take ever more extreme measures to stand out.

For instance, parliamentary reviews of the 2009 'Climategate' scandal at the University of East Anglia in Norwich, UK, reported evidence of scientific misconduct (see go.nature.com/d6bdco). The allegations included questionable journal refereeing to promote a particular scientific line (see also *Nature* <http://doi.org/ftb9hc>; 2010). Instead, journals should be supported as places where unsettled science is refined by open debate. But, compared with 30 years ago, they do seem less willing to publish negative results or cautionary reviews that temper unbridled enthusiasm — perhaps because of ratings wars.

In another example, a May 2011 article in *Times Higher Education* reported on the cover-up of data that would otherwise have prevented a cancer drug from entering phase III clinical trials, raising and dashing patients' hopes and putting lives at risk on a false premise.

There is also a tendency to issue breathless press releases to accompany publication of even modest advances, and for entries in the 'future impact' section of grant application forms to be loaded with ludicrous hyperbole.

It is any wonder that trust in scientists is starting to decline (see go.nature.com/3xdcoy)?

Michael Kelly *University of Cambridge, UK.* mjk1@cam.ac.uk

Don't do big-data science backwards

Large open-access data sets offer unprecedented opportunities for scientific discovery — the current global collapse of bee and frog populations are classic examples. However, we must resist the temptation to do science backwards by posing questions after, rather than before, data analysis.

A scant understanding of the context in which data sets were collected can lead to poorly framed questions and results, and to conclusions that are plain wrong. Scientists intending to make use of large composite data sets need to work closely with those responsible for gathering the data. Standard scientific principles and practice then demand that they first frame the important questions, then design and execute the data analyses needed to answer them.

David B. Lindenmayer *Australian National University, Canberra, Australia.* david.lindenmayer@anu.edu.au
Gene E. Likens *University of Connecticut, Storrs, USA.*

CONTRIBUTIONS

Correspondence may be sent to correspondence@nature.com after consulting the author guidelines at go.nature.com/cmchno. Alternatively, readers may comment online: www.nature.com/nature.

Andreev states taken to the next level

The Andreev bound states are a hallmark of the theory of superconducting weak links. Using the versatile Josephson effect as source, device and detector, these spin-1/2 states have been directly observed. [SEE LETTER P.312](#)

SIMON GUSTAVSSON & WILLIAM D. OLIVER

The Josephson junction is a remarkably versatile superconducting device. Apply an electrical current to the junction, and it becomes a tunable inductor that can pass a dissipationless supercurrent. Apply a voltage to it, and the junction becomes a source of electromagnetic radiation. Collectively known as the Josephson effect, these and related electronic-transport phenomena have been observed across a variety of thin non-superconducting barriers called weak links, including insulating and metallic tunnel junctions, nanowires and molecules. Although the Josephson effect was predicted more than 50 years ago¹, it took until the early 1990s for researchers to develop fully a mesoscopic transport theory that could treat these various weak links in a unified manner. On page 312 of this issue, Bretheau *et al.*² provide direct spectroscopic evidence for a key tenet of this theory — a degree of freedom internal

to the weak links called the Andreev bound-state doublet.

When a material is cooled below the critical temperature at which it becomes superconducting, its electrons pair up to form Cooper pairs. These pairs have a quantum-mechanical affinity for one another that encourages them to condense into a common, quantum-mechanical state, which is described by a wavefunction with a well-defined phase. This process opens up an energy gap in the material's spectrum of allowed energy states, preventing low-energy electronic excitations and thereby allowing electrical current to flow without resistance.

When two such superconductors are brought into close proximity, the Cooper pairs can even transit from one superconductor to the other. An example is the superconductor–insulator–superconductor tunnel junction (Fig. 1). In essence, when the superconductors are close enough that their quantum-mechanical wavefunctions coherently extend through

the thin insulator and into one another, then the wavefunctions interfere such that the phases twist across the insulator and electrical current can flow. The wave-mechanics description of Josephson junctions given above intuitively describes their supercurrent³, but what is the internal mechanism that mediates this transport? In other words, what is going on in between the superconductors?

To answer this question, we need a corresponding ‘particle picture’ for the Josephson effect. In this picture, transport between the two superconductors is realized by particle-scattering processes within an array of weak links, each with a unique transmission probability and capable of accommodating a single Cooper pair in transit⁴. Each weak link has two discrete-energy states — an Andreev bound-state doublet — with energies related to both the phase difference across the weak link and the transmission probability through it. The bound-state energies generally lie within the superconducting energy gap. The ground

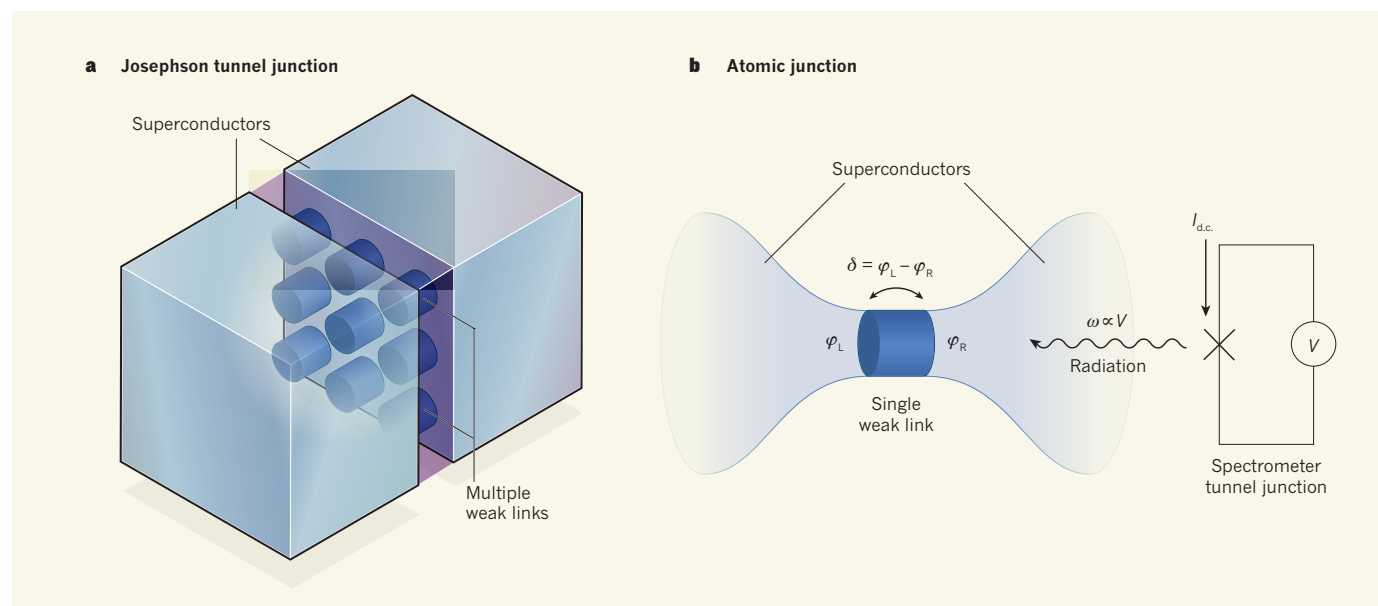


Figure 1 | Photon-absorption spectroscopy of an Andreev bound-state doublet. **a**, In a typical macroscopic Josephson tunnel junction, the two superconductors are connected by many weak links in parallel through a thin slab of metallic or insulating material. **b**, Using an atomic junction, one can access a single weak link. The link is characterized by the phase difference δ between the two superconductors, which determines the energies of the link's Andreev bound states; ϕ_L and ϕ_R denote the phases of the two

superconductors. Bretheau *et al.*² probed the link with radiation from a spectrometer tunnel junction (denoted by the cross) subject to an applied voltage (V). The spectrometer's current (I_{dc}) peaks whenever the frequency of the radiation (ω) from the junction, which is proportional to V , matches the energy separation of the Andreev bound states. This allows the measurement of the energy-level spectrum for the Andreev bound-state doublet.

state of the doublet accommodates current in the 'forward' direction (the supercurrent usually observed in junctions), whereas the excited state corresponds to current in the 'reverse' direction. Importantly, the Andreev bound-state doublet is a system of spin-1/2 that is localized to the weak link and behaves as an internal degree of freedom associated with the Josephson effect. It is this internal feature that Bretheau *et al.* detected directly using the technique of photon-absorption spectroscopy.

The authors used several aspects of the Josephson effect to carry out photon-absorption spectroscopy of the Andreev doublet. First, to access the doublet from a single weak link, they used a mechanically controllable atomic break junction (a narrow bridge containing only a few atoms) formed on a flexible substrate⁵. By applying a mechanical force *in situ* to the cryogenically cooled break junction, the researchers could reduce the number of channels down to as low as a single weak link with a particular transmission probability. Second, to change the phase difference across the atomic junction and thereby tune the Andreev-doublet energies, Bretheau *et al.* fabricated the break junction as part of a superconducting loop containing a much larger Josephson tunnel junction, thereby forming a sensitive magnetometer known as a superconducting quantum interference device. In this configuration, a magnetic field applied to the loop sets the phase difference across the atomic junction and makes the bound-state energies tunable.

Third, to drive transitions between the ground and excited Andreev bound states, the authors used a separate tunnel junction as a spectrometer (Fig. 1). Applying a voltage to the spectrometer junction causes it to emit electromagnetic radiation at a frequency that is proportional to the applied voltage and is tunable over a wide range — up to around 80 gigahertz. Finally, owing to conservation of energy, a unidirectional (d.c.) current flows through the spectrometer junction that is proportional to the absorption rate of the radiation emitted by the atomic junction or the surrounding environment. Therefore, determining the spectrometer's d.c. current is a measurement of the absorption rate. The spectrometer Josephson junction thus serves both as a microwave generator and a detector in this experiment. With this set-up, Bretheau *et al.* carried out photon-absorption spectroscopy of the Andreev bound-state doublet, observing the expected trend in transition frequency as a function of the phase difference across the atomic junction.

It should be noted that in several earlier experimental works (for example, ref. 6), the presence of Andreev bound states was required to explain the observed phenomena, serving as indirect evidence for their existence. More recently, direct evidence was provided by applying a technique called tunnelling spectroscopy

to a carbon nanotube weak link⁷ and to a graphene quantum dot⁸ — complementary studies that observed alternative microscopic configurations of the bound states.

Bretheau and colleagues' experiment stands as the first photon-absorption spectroscopy of the spin-1/2 Andreev bound-state doublet. These doublets could be used as a quantum logic element or, in conjunction with interactions between the electron's spin and orbital degrees of freedom, to realize a Majorana state — an elusive spin-1/2 particle that is its own antiparticle. This experimental technique may also shed light on the perennial problem of quasi-particle poisoning in coherent superconducting devices. Indeed, by driving the bound-state transition, this work opens up experiments to a new level of Andreev physics. ■

BIOGEOCHEMISTRY

Carbon dioxide and water use in forests

Plants are expected to respond to rising levels of atmospheric carbon dioxide by using water more efficiently. Direct evidence of this has been obtained from forests, but the size of the effect will prompt debate. SEE LETTER P.324

BELINDA MEDLYN & MARTIN DE KAUWE

In a study published on page 324 of this issue, Keenan *et al.*¹ report that the efficiency with which forests use water has increased over the past 20 years, and conclude that this is a consequence of rises in the concentration of atmospheric carbon dioxide.* The findings call for a reassessment of models of the terrestrial carbon cycle.

The concentration of CO₂ in the atmosphere is rising at an unprecedented rate. In May this year, it reached 400 parts per million, 43% above the pre-industrial concentration of 280 p.p.m. (ref. 2). Much of this increase has occurred in recent decades, with the rate of increase over the past 20 years being 5% per decade². This drastic upsurge in atmospheric CO₂ should have stimulated plant productivity worldwide, because we know from experiments that rising CO₂ concentrations increase the rate of photosynthesis and reduce water use in plants³. Such effects are fundamental to our current understanding of the carbon cycle — for example, most terrestrial carbon-cycle models explain the current land sink for carbon by assuming that rising CO₂ levels have enhanced plant productivity⁴.

However, detecting the effects of rising

*This article and the paper under discussion¹ were published online on 10 July 2013.

Simon Gustavsson is in the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139–4307, USA. **William D. Oliver** is in the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02420–9108, USA. e-mails: simongus@mit.edu; oliver@ll.mit.edu

1. Josephson, B. D. *Phys. Lett.* **1**, 251 (1962).
2. Bretheau, L. *et al. Nature* **499**, 312–315 (2013).
3. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* (Basic Books, 2011).
4. Hekking, F. W. J., Schon, G. & Averin, D. V. *Proc. NATO Adv. Res. Workshop Mesoscopic Superconductivity* (NATO-ASI, 1994).
5. van Ruitenbeek, J. M. *et al. Rev. Sci. Instrum.* **67**, 108–111 (1996).
6. Fueschle, M. *et al. Phys. Rev. Lett.* **102**, 127001 (2009).
7. Pillet, J.-D. *et al. Nature Phys.* **6**, 965–969 (2010).
8. Dirks, T. *et al. Nature Phys.* **7**, 386–390 (2011).

CO₂ concentrations on terrestrial vegetation outside controlled experiments has proven remarkably difficult, provoking numerous debates about whether such effects are really occurring^{5–10}. There are few high-quality, long-term records of plant productivity and water use that can be used to test for such effects. The main types of data come from plot surveys, tree-ring records, satellite images, aerial photographs and measurements of stream flow. Each of these is an indirect measurement and has a relatively coarse time resolution. Even where trends in these data have been detected, it has been extremely difficult to attribute them to rising CO₂ levels, because simultaneous changes in many confounding factors — such as rainfall, temperature, land use and fire frequency — have occurred¹¹.

Keenan *et al.* bring a new source of data to bear on this problem. The eddy-covariance technique, developed in the 1980s to quantify the exchange of gases between the atmosphere and land areas, has revolutionized plant-ecosystem science because it continuously monitors the functioning of whole ecosystems on an hourly timescale¹². Using instruments mounted above a vegetation canopy, eddy covariance can be used to measure the carbon uptake and water use of whole ecosystems on a spatial scale of up to one square kilometre. Over the past 20 years, eddy-covariance towers

9. Clark, D. A., Clark, D. B. & Oberbauer, S. F. *J. Geophys. Res.* **118**, 1–12 (2013).
10. Silva, L. C. R. & Anand, M. *Glob. Ecol. Biogeogr.* **22**, 83–92 (2013).
11. Donohue, R. J., McVicar, T. R. & Roderick, M. L. *Glob. Change Biol.* **15**, 1025–1039 (2009).
12. Baldocchi, D. *Aust. J. Bot.* **56**, 1–26 (2008).
13. Wong, S. C., Cowan, I. R. & Farquhar, G. D. *Plant Physiol.* **78**, 821–825 (1985).
14. Drake, B. G., González-Meler, M. A. & Long, S. P. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **28**, 609–639 (1996).
15. Ainsworth, E. A. & Rogers, A. *Plant Cell Environ.* **30**, 258–270 (2007).
16. Barton, C. V. M. *et al. Glob. Change Biol.* **18**, 585–595 (2012).
17. De Kauwe, M. G. *et al. Glob. Change Biol.* **19**, 1759–1779 (2013).

RNA BIOPHYSICS

A three-state balancing act

How do pathogens survive temperature variations? At a molecular level, one bacterial species seems to regulate gene expression in response to temperature through structural equilibria in corresponding RNA sequences. [SEE LETTER P.355](#)

RONALD MICURA

In bacteria, many messenger RNA molecules carry a regulatory segment called a riboswitch. Specific binding of small ligand molecules, such as adenine, to this segment determine whether the riboswitch mRNA will be translated into a protein^{1–4}. On page 355 of this issue, Reining *et al.*⁵ show that regulation of gene expression through such ‘riboswitching’ is coupled to temperature sensing. The authors investigate the adenine-sensitive riboswitch from the pathogenic bacterium *Vibrio vulnificus*⁶, and demonstrate that efficient RNA regulation at different temperatures — those of the bacterium’s marine habitat and its human host — requires a change in the riboswitch’s structural behaviour from a two-state pattern to a three-state one.*

A typical riboswitch consists of two domains at the 5' end of the mRNA: a ligand-binding aptamer and an adjoining expression platform, which can have one of two mutually exclusive structures depending on whether the aptamer is in the ligand-bound or ligand-unbound state. The structural change in the expression platform signals that gene expression should be turned on or off.

During gene transcription, a polymerase enzyme synthesizes first the aptamer and then the expression platform⁷. The sequential release of riboswitch domains from the polymerase is especially meaningful for transcription-controlling riboswitches, which act under kinetic control. To direct the folding of the expression platform, the aptamer domain of the growing RNA chain must bind rapidly to its ligand, which requires high ligand concentrations⁸; otherwise, the resulting full-length mRNA becomes trapped in a default fold that cannot respond to the ligand⁹. In

transcription-controlling riboswitches, the two mutually exclusive structures are generally referred to as terminator and antiterminator folds, causing cessation of polymerase activity and continuation of mRNA synthesis, respectively.

Bacterial transcription is tightly coupled to translation. In translation-controlling riboswitches, the molecular mechanism relies on either sequestration or liberation of the Shine–Dalgarno sequence — the mRNA site that binds to cellular organelles known as ribosomes to initiate translation. In contrast to transcription-controlling riboswitches, most translation-controlling riboswitches act under thermodynamic control⁹. Consequently, ligand-dependent control of translation is maintained even for a full-length mRNA. For these riboswitches, therefore, two states (ligand-bound and -unbound) seem sufficient to turn translation on and off.

Reining and colleagues have found that, for robust functioning, their translation-controlling adenine-sensing riboswitch must occur in three structural states. The authors investigated the full-length (more than 100-nucleotide) riboswitch domain at single-nucleotide resolution. They also determined a complete set of thermodynamic and kinetic parameters for folding and ligand-binding of this RNA, under conditions that involved varying the concentrations of RNA, magnesium ions (a factor mediating structure formation) and adenine, as well as, importantly, temperature.

The researchers find that a ligand-free (apo) form of the riboswitch exists in a pre-equilibrium of two structurally distinct aptamer folds (Fig. 1). One of the structures (apoA) can bind to the ligand, exhibiting a structure that resembles a third structural state — the adenine-bound holo form. The other structure (apoB) adopts a different fold and cannot interact with the ligand. But why does



50 Years Ago

A technique has been used with limited success to obtain simultaneous measurement of cosmic radiation at two different altitudes. The method consists of suspending two packets of nuclear emulsion plates from the one balloon, while maintaining a constant vertical separation of 10,000–27,000 ft. between the packets ... in this way the two packets do not separate in latitude and longitude—a factor which enters if two separate independent balloons are flown.

This has been achieved by carrying aloft a cone of nylon string which was allowed to unwind at a predetermined height, leaving one packet of plates suspended at the balloon while the other fell at the end of the string ... As the alarm rings, the alarm winder releases the key ... holding the lower packet and target, which then fall away freely, unwinding the string in the process. Twenty thousand feet of string unwinds in approximately 10 min. **From *Nature* 20 July 1963**

100 Years Ago

The Potato: A Compilation of Information from Every Available Source. By E. H. Grubb and W. S. Guilford.

There are men who, having attained to wealth and fame by the agency of some humble instrument, basely repudiate and kick over the ladder by which they have risen. Not so the authors of the first book on our list. The potato has “made” them, and in return they proceed to “make” the potato ... the authors are so evidently enthusiastic, and discourse so eloquently on the merits of their subject, that we are carried along with them, and forget that, after all, they are only talking about potatoes, and not about alpine plants or roses.

From *Nature* 17 July 1913

*This article and the paper under discussion⁵ were published online on 10 July 2013.

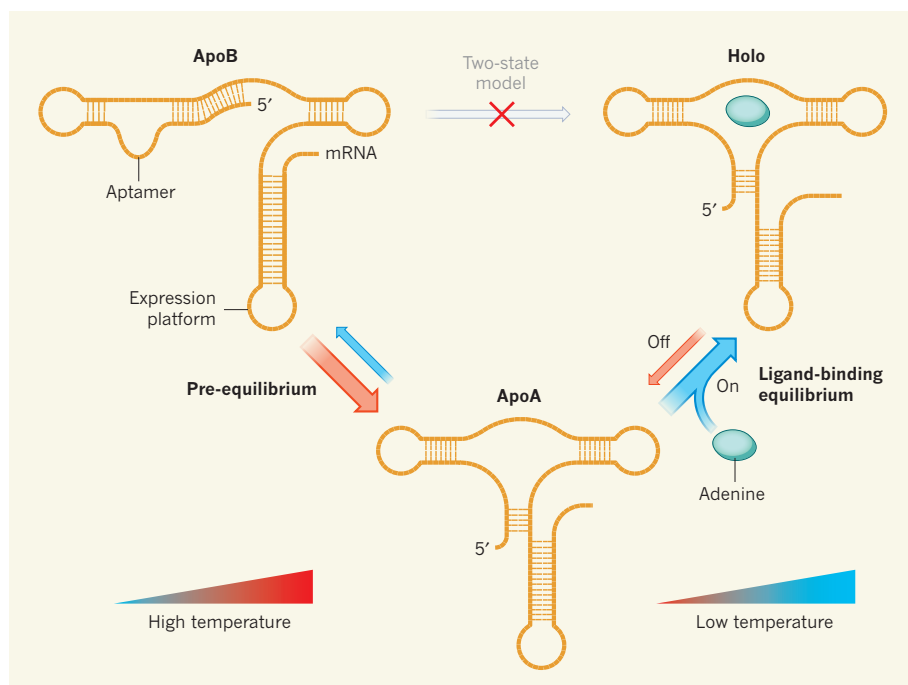


Figure 1 | Three states of a riboswitch. Riboswitches consist of an aptamer and an expression platform. Reining *et al.*⁵ find that, in contrast to the accepted two-state model, a three-state model exists for the structural behaviour of an adenine-sensing bacterial riboswitch. They show that the ligand-free apo form of this riboswitch can exist in two pre-equilibrium states: an apoA state that can bind a ligand (adenine), and an apoB state that cannot. On binding adenine, the apoA structure adopts a third (holo) state. The structural pre-equilibrium between apoA and apoB states counterbalances the temperature-dependent equilibrium between the apoA and holo states, guaranteeing effective regulation, over a broad temperature range, of the gene corresponding to the RNA sequence that carries the riboswitch.

this bistable sequence element exist?

Reining *et al.* provide a clear answer: this pre-equilibrium is markedly temperature dependent and counterbalances the temperature-dependent changes in ligand affinity. At low temperatures, when adenine has a high affinity for the RNA, only a small population of the adenine-binding apoA form is available, but it is sufficient to achieve high riboswitching efficiency (Fig. 1). At higher temperatures, this population is significantly increased and compensates for the lower RNA affinity of the ligand. Thus, whereas in a pure two-state model switching efficiency would vary drastically with environmental changes, in this three-state model the switching efficiency remains robust.

The authors further perform a coupled transcription–translation assay to underline the significance of their biophysical findings. They show that a mutant riboswitch that cannot adopt the apoB structure and exists only in the binding-competent apoA conformation lacks sensitivity to changes in adenine concentration and so cannot control gene expression in response to varying ligand doses. This mutant is also insensitive to changes in temperature and is always in an ‘on’ state.

Nature probably contains further examples of three-state structural behaviour for RNA that balances RNA-regulated gene expression. Bistable elements are widespread in RNA

sequences and have been thoroughly characterized in at least one other riboswitch system — a preQ₁ class I riboswitch from the human pathogen *Fusobacterium nucleatum*¹⁰. Moreover, temperature-sensitive RNAs that control gene expression, ‘RNA thermometers’¹¹, are prevalent. Coupling ligand and temperature sensitivities seems logical, although the three-state mechanism discovered by Reining *et al.*

PARTICLE PHYSICS

Let it B

A phenomenon known as CP asymmetry, which explains our very existence, has been observed in the decays of B_s^0 mesonic particles. The finding represents yet another triumph of the standard model of particle physics.

YOSEF NIR

The first observation of matter–antimatter asymmetry in the decays of particles known as B_s^0 mesons has been reported in *Physical Review Letters* by the LHCb collaboration¹ (Aaij *et al.*) at the Large Hadron Collider at CERN, near Geneva, Switzerland. This measurement reinforces the standard model of particle physics and, in particular, its

compensates for temperature fluctuations rather than using this parameter to directly trigger a gene response.

As the present paper shows, mechanistic insights into riboswitch function can provide a deeper understanding of the potential adaptive strategies used by bacteria in different environments. Such insights may also be of fundamental interest for exploring the possibility of targeting riboswitch RNAs with antibacterial drugs — for example, by interfering with the delicate balance of the RNA-structure equilibria. Furthermore, these findings will influence the design of RNA biosensors, which represent emerging tools for live-cell imaging¹². They will also affect the use of riboswitches in synthetic-biology applications for reprogramming cells to autonomously perform complex tasks, such as ‘seek-and-destroy’ herbicides¹³. ■

Ronald Micura is at the Institute of Organic Chemistry and the Center for Molecular Biosciences, Leopold Franzens University, 6020 Innsbruck, Austria.
e-mail: ronald.micura@uibk.ac.at

- Breaker, R. R. *Mol. Cell* **43**, 867–879 (2011).
- Serganov, A. & Patel, D. J. *Annu. Rev. Biophys. Biomol. Struct.* **41**, 343–370 (2012).
- Serganov, A. & Nudler, E. *Cell* **152**, 17–24 (2013).
- Liberman, J. A. & Wedekind, J. E. *Wiley Interdisc. Rev. RNA* **3**, 369–384 (2012).
- Reining, A. *et al. Nature* **499**, 355–359 (2013).
- Lemay, J.-F. *et al. PLoS Genet.* **7**, e1001278 (2011).
- Garst, A. D., Edwards, A. L. & Batey, R. T. *Cold Spring Harb. Perspect. Biol.* **3**, a003533 (2011).
- Al-Hashimi, H. M. & Walter, N. G. *Curr. Opin. Struct. Biol.* **18**, 321–329 (2008).
- Zhang, J., Lau, M. W. & Ferré-D’Amaré, A. R. *Biochemistry* **49**, 9123–9131 (2010).
- Rieder, U., Kreutz, C. & Micura, R. *Proc. Natl Acad. Sci. USA* **107**, 10804–10809 (2010).
- Kortmann, J. & Narberhaus, F. *Nature Rev. Microbiol.* **10**, 255–365 (2012).
- Kellenberger, C. A., Wilson, S. C., Sales-Lee, J. & Hammond, M. C. *J. Am. Chem. Soc.* **135**, 4906–4909 (2013).
- Sinha, J., Reyes, S. J. & Gallivan, J. P. *Nature Chem. Biol.* **6**, 464–470 (2010).

explanation of how the weak force, which governs radioactive decays, distinguishes matter from antimatter.

Matter particles and their antiparticles have opposite charge (C) and parity (P), the latter meaning that, under spatial-inversion transformation, the direction of a particle’s spin relative to its direction of motion is reversed. The difference in the laws of physics obeyed by matter and antimatter is experimentally

explored by comparing the rate of a process — such as the rate at which a particle will decay into other, lighter particles — with the rate of its CP-related process, in which all the initial and final particles are replaced with the corresponding antiparticles. If these rates are measured to be different, then a ‘CP asymmetry’ has been observed. Measurements of CP asymmetries are interesting for two reasons. First, they provide stringent tests of the standard model of particle physics. Second, they may provide hints that could point scientists in the direction of a solution to the mystery of the Universe’s matter–antimatter imbalance.

Among the four known forces of nature — the standard model’s weak, strong and electromagnetic forces, and the gravitational force — the weak force is the only one that distinguishes between matter and antimatter. But even the weak force would not do so if it were not that three quark particles exist for each of two charge types: the down, strange and bottom quarks, which have a charge that is minus one-third of the proton’s charge; and the up, charm and top quarks, which have a charge that is two-thirds of the proton’s charge. With three quarks of either charge, the weak force has a single coupling constant (a parameter that quantifies the force’s strength), which is different for particles and antiparticles^{2,3}. If, as the standard model predicts, all CP asymmetries are proportional to this single coupling constant, then the sizes of these asymmetries are correlated. By contrast, most theories that go beyond the standard model have many independent coupling constants that distinguish particles from antiparticles, and thus predict deviations from the standard-model correlations.

Mesons are bound states of one quark and one antiquark. Until the LHCb measurement was made, CP asymmetries had been observed in the decays of three types of meson: K^0 mesons, B^0 mesons, and B^\pm mesons. The LHCb experiment is the first to measure CP asymmetry in the decay of a fourth type of meson, the B_s^0 meson. The fact that, within the standard model, the difference in the laws of physics followed by matter and antimatter is encoded in a single parameter provides a particularly strong correlation between this asymmetry and an asymmetry in B^0 decay that has been measured by several experiments with high accuracy. Although neither of these asymmetries, nor the corresponding decay rates, can be theoretically predicted with any accuracy, the product of the CP asymmetry with the decay rate should, according to the standard model, be equal (to a good approximation) between the B_s^0 decay and the B^0 decay^{4,5}. The LHCb measurement is consistent with the predicted equality at the level of a few per cent, implying yet another triumph of the standard model.

In many extensions of the standard model, new particles are predicted to interact more strongly with the heavier quarks (strange, bottom, charm and top) than with the lighter ones

(up and down). The K^0 mesons, the B^0 mesons and the B^\pm mesons all have either a down (anti)quark or an up (anti)quark. But the B_s^0 mesons are different: they are made up of a strange quark and a bottom antiquark. Their antiparticles are composed of a bottom quark and a strange antiquark. Thus, scientists had hoped that although the effects of new physics were negligibly small in all previously measured CP asymmetries, they would be large enough to be observed in B_s^0 decays. Frustratingly, this seems not to be the case.

“The LHCb measurement reinforces the standard model’s explanation of how the weak force distinguishes matter from antimatter.”

The standard-model prediction² of how the weak force distinguishes between matter and antimatter therefore continues to successfully describe all measurements of CP asymmetries in meson decays. However, the standard model fails to explain the Universe’s matter–antimatter imbalance. All structures in the Universe, from clusters of galaxies to human cells, are made of matter: protons, neutrons and electrons. Their antiparticles — the antiprotons, antineutrons and positrons — are not found in the Universe at large. If the laws of nature were identical for matter and antimatter, then particles and antiparticles would have been created in equal amounts and would then have annihilated each other,

leaving only pure radiation and no matter structures in the Universe. Our very existence is possible only because of CP asymmetries. The standard model allows all antimatter to disappear from the Universe, but it predicts that the amount of surviving matter is many orders of magnitude smaller than observed.

Therefore, there must exist a force, as yet unknown to us, that distinguishes matter from antimatter in a way that is much stronger than that of the weak force. Theorists have made various suggestions as to what this new force might be. To disclose the nature of this force, further hints from experiments are needed. Searching for CP asymmetries in neutrino ‘oscillations’ of one type into another and for the electric-dipole moments of the neutron and the electron, in addition to measuring CP asymmetries in B_s^0 -meson decays, seem the most promising avenues through which to obtain such hints. The consistency of the LHCb measurement with the standard model provides further motivation to pursue the other searches even more vigorously. ■

Yosef Nir is in the Department of Particle Physics, Weizmann Institute of Science, Rehovot 76100, Israel.
e-mail: yosef.nir@weizmann.ac.il

1. Aaij, R. *et al.* *Phys. Rev. Lett.* **110**, 221601 (2013).
2. Kobayashi, M. & Maskawa, T. *Prog. Theor. Phys.* **49**, 652–657 (1973).
3. Jarlskog, C. *Phys. Rev. Lett.* **55**, 1039–1042 (1985).
4. Lipkin, H. J. *Phys. Lett. B* **621**, 126–132 (2005).
5. Gronau, M. & Rosner, J. L. *Phys. Lett. B* **482**, 71–76 (2000).

CANCER

Calculated treatment

Mathematical modelling linked with patient data suggests that combination therapy is more effective than sequential treatment at preventing drug resistance in cancer. This predictive approach may pave the way for personalized therapies.

**NATALIA L. KOMAROVA
& C. RICHARD BOLAND**

In a study published in *eLIFE*, Bozic *et al.*¹ use a mathematical approach to examine tumour evolution and response to chemotherapy. In one example, they describe a patient who had the skin cancer melanoma characterized by an estimated tumour burden of 9.8×10^{10} cells and 8 metastatic lesions. Their modelling predicts a 0% chance of disease control using a single drug, but that the likelihood of successful treatment could rise to 88% during combined therapy with two drugs. The approach offers a brave, quantitative look at designing targeted therapy for cancer.

The search for cancer treatments has traversed a long and thorny path, with more

failures and disappointments than glimpses of success. A major breakthrough was achieved with the development of a drug called imatinib in the 1990s. This inhibitor of tyrosine kinase enzymes showed breathtaking success for treating chronic myelogenous leukaemia (CML). Imatinib and other small-molecule inhibitors ‘recognize’ and attack cancer cells, but spare normal cells, thereby reducing side effects compared with conventional chemotherapy. Since the discovery of imatinib, dozens of other such inhibitors have been developed for treating different cancers. However, the initial excitement surrounding these drugs was tempered by the appearance of drug resistance — the phenomenon in which disease returns a few months after initial treatment success².

explored by comparing the rate of a process — such as the rate at which a particle will decay into other, lighter particles — with the rate of its CP-related process, in which all the initial and final particles are replaced with the corresponding antiparticles. If these rates are measured to be different, then a ‘CP asymmetry’ has been observed. Measurements of CP asymmetries are interesting for two reasons. First, they provide stringent tests of the standard model of particle physics. Second, they may provide hints that could point scientists in the direction of a solution to the mystery of the Universe’s matter–antimatter imbalance.

Among the four known forces of nature — the standard model’s weak, strong and electromagnetic forces, and the gravitational force — the weak force is the only one that distinguishes between matter and antimatter. But even the weak force would not do so if it were not that three quark particles exist for each of two charge types: the down, strange and bottom quarks, which have a charge that is minus one-third of the proton’s charge; and the up, charm and top quarks, which have a charge that is two-thirds of the proton’s charge. With three quarks of either charge, the weak force has a single coupling constant (a parameter that quantifies the force’s strength), which is different for particles and antiparticles^{2,3}. If, as the standard model predicts, all CP asymmetries are proportional to this single coupling constant, then the sizes of these asymmetries are correlated. By contrast, most theories that go beyond the standard model have many independent coupling constants that distinguish particles from antiparticles, and thus predict deviations from the standard-model correlations.

Mesons are bound states of one quark and one antiquark. Until the LHCb measurement was made, CP asymmetries had been observed in the decays of three types of meson: K^0 mesons, B^0 mesons, and B^\pm mesons. The LHCb experiment is the first to measure CP asymmetry in the decay of a fourth type of meson, the B_s^0 meson. The fact that, within the standard model, the difference in the laws of physics followed by matter and antimatter is encoded in a single parameter provides a particularly strong correlation between this asymmetry and an asymmetry in B^0 decay that has been measured by several experiments with high accuracy. Although neither of these asymmetries, nor the corresponding decay rates, can be theoretically predicted with any accuracy, the product of the CP asymmetry with the decay rate should, according to the standard model, be equal (to a good approximation) between the B_s^0 decay and the B^0 decay^{4,5}. The LHCb measurement is consistent with the predicted equality at the level of a few per cent, implying yet another triumph of the standard model.

In many extensions of the standard model, new particles are predicted to interact more strongly with the heavier quarks (strange, bottom, charm and top) than with the lighter ones

(up and down). The K^0 mesons, the B^0 mesons and the B^\pm mesons all have either a down (anti)quark or an up (anti)quark. But the B_s^0 mesons are different: they are made up of a strange quark and a bottom antiquark. Their antiparticles are composed of a bottom quark and a strange antiquark. Thus, scientists had hoped that although the effects of new physics were negligibly small in all previously measured CP asymmetries, they would be large enough to be observed in B_s^0 decays. Frustratingly, this seems not to be the case.

“The LHCb measurement reinforces the standard model’s explanation of how the weak force distinguishes matter from antimatter.”

The standard-model prediction² of how the weak force distinguishes between matter and antimatter therefore continues to successfully describe all measurements of CP asymmetries in meson decays. However, the standard model fails to explain the Universe’s matter–antimatter imbalance. All structures in the Universe, from clusters of galaxies to human cells, are made of matter: protons, neutrons and electrons. Their antiparticles — the antiprotons, antineutrons and positrons — are not found in the Universe at large. If the laws of nature were identical for matter and antimatter, then particles and antiparticles would have been created in equal amounts and would then have annihilated each other,

leaving only pure radiation and no matter structures in the Universe. Our very existence is possible only because of CP asymmetries. The standard model allows all antimatter to disappear from the Universe, but it predicts that the amount of surviving matter is many orders of magnitude smaller than observed.

Therefore, there must exist a force, as yet unknown to us, that distinguishes matter from antimatter in a way that is much stronger than that of the weak force. Theorists have made various suggestions as to what this new force might be. To disclose the nature of this force, further hints from experiments are needed. Searching for CP asymmetries in neutrino ‘oscillations’ of one type into another and for the electric-dipole moments of the neutron and the electron, in addition to measuring CP asymmetries in B_s^0 -meson decays, seem the most promising avenues through which to obtain such hints. The consistency of the LHCb measurement with the standard model provides further motivation to pursue the other searches even more vigorously. ■

Yosef Nir is in the Department of Particle Physics, Weizmann Institute of Science, Rehovot 76100, Israel.
e-mail: yosef.nir@weizmann.ac.il

1. Aaij, R. *et al.* *Phys. Rev. Lett.* **110**, 221601 (2013).
2. Kobayashi, M. & Maskawa, T. *Prog. Theor. Phys.* **49**, 652–657 (1973).
3. Jarlskog, C. *Phys. Rev. Lett.* **55**, 1039–1042 (1985).
4. Lipkin, H. J. *Phys. Lett. B* **621**, 126–132 (2005).
5. Gronau, M. & Rosner, J. L. *Phys. Lett. B* **482**, 71–76 (2000).

CANCER

Calculated treatment

Mathematical modelling linked with patient data suggests that combination therapy is more effective than sequential treatment at preventing drug resistance in cancer. This predictive approach may pave the way for personalized therapies.

NATALIA L. KOMAROVA
& C. RICHARD BOLAND

In a study published in *eLIFE*, Bozic *et al.*¹ use a mathematical approach to examine tumour evolution and response to chemotherapy. In one example, they describe a patient who had the skin cancer melanoma characterized by an estimated tumour burden of 9.8×10^{10} cells and 8 metastatic lesions. Their modelling predicts a 0% chance of disease control using a single drug, but that the likelihood of successful treatment could rise to 88% during combined therapy with two drugs. The approach offers a brave, quantitative look at designing targeted therapy for cancer.

The search for cancer treatments has traversed a long and thorny path, with more

failures and disappointments than glimpses of success. A major breakthrough was achieved with the development of a drug called imatinib in the 1990s. This inhibitor of tyrosine kinase enzymes showed breathtaking success for treating chronic myelogenous leukaemia (CML). Imatinib and other small-molecule inhibitors ‘recognize’ and attack cancer cells, but spare normal cells, thereby reducing side effects compared with conventional chemotherapy. Since the discovery of imatinib, dozens of other such inhibitors have been developed for treating different cancers. However, the initial excitement surrounding these drugs was tempered by the appearance of drug resistance — the phenomenon in which disease returns a few months after initial treatment success².

Cancer is a process of Darwinian evolution played out in a particular organ. Normal cells divide and die, and each division brings a small chance of genetic change. Most such mutations are deleterious, and the cells die without leaving offspring, but some confer new properties that promote growth or survival and can lead to cancer. The greatest challenge for drugs targeted at such cells is the further evolution of mutations that confer drug resistance. Combining multiple drugs that have distinct mechanisms of action might provide a solution to this problem. The concept of a 'drug cocktail' was introduced in 1996 in the context of treating AIDS. There, the emergence of viral strains resistant to single drug treatments rendered all previous attempts to control the disease unsuccessful. Viral evolution in a patient is conceptually similar to the evolution of cancer cells, so similar treatment strategies might work well for both diseases.

Current common practice for treating cancers with small-molecule inhibitors is to administer the agents sequentially, starting with a 'first-line' drug and switching to 'second-line' therapies if the tumour relapses. Bozic *et al.* assessed the effectiveness of this approach using sophisticated mathematical techniques and data from patients with melanoma or with pancreatic or colorectal cancers. They convincingly demonstrate that a sequential strategy "precludes any chance for cure", even in the best-case scenario in which no single mutants confer resistance to both drugs. However, they show that simultaneously combining two or more drugs can provide much-needed hope for patients.

As demonstrated last year³, drug-resistant mutants typically exist at low levels in tumours before the beginning of treatment. Treatment with a single drug gives a competitive advantage to mutants resistant to that drug, such that by the time of the switch to a second-line therapy, there is a high chance that a mutant that is also resistant to the second drug (a doubly resistant mutant) has already emerged (Fig. 1). But combination therapy eliminates cells that are singly resistant to either drug and therefore — because the likelihood of a doubly mutated cell emerging in such a population is low — greatly increases the chance of success.

The greatest obstacle for combination treatments is the phenomenon of cross-resistance, in which a single mutation confers resistance to more than one drug. But even if such mutants are generated, the authors estimate that combination treatment can be beneficial in some cases, whereas single-drug and sequential strategies offer no hope.

Previous mathematical analyses have also shown⁴ that cyclic treatments are ineffective compared with combination treatments, and have led to the proposal⁵ that a combination of three anticancer drugs would be needed to treat CML. It has also already been argued⁶ that, even in the presence of cross-resistant

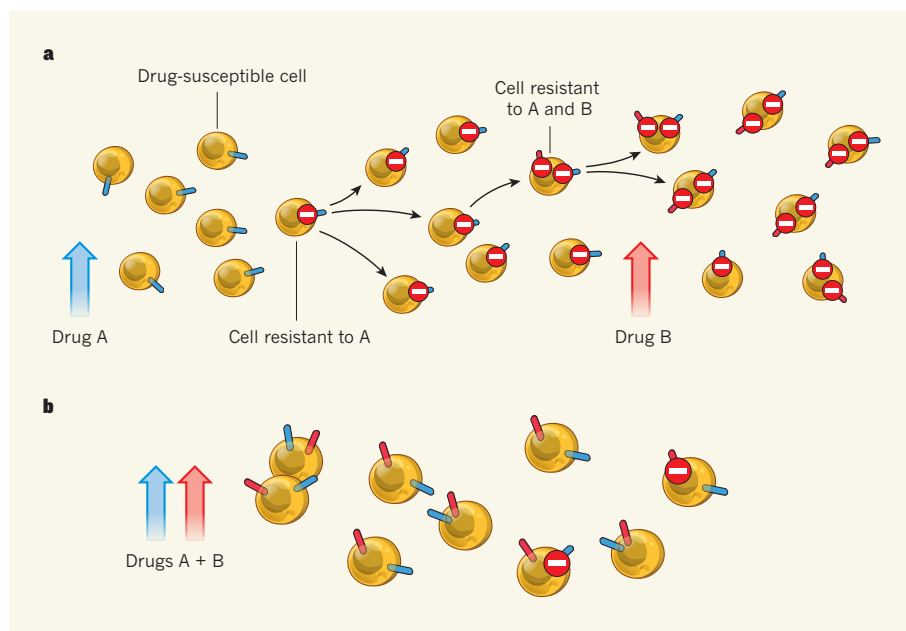


Figure 1 | Single-drug versus combination therapy. **a**, During therapy with one drug alone, a cell that acquires a mutation that confers resistance to the drug will be at a proliferative advantage. By the time this is recognized and treatment with a second drug is started, it is likely that a cell resistant to both drugs will already have emerged. **b**, Starting therapy with both drugs simultaneously means that cells acquiring single resistance will be immediately eliminated by the other drug. Bozic *et al.*¹ use mathematical modelling to show that this approach increases the chance of effective treatment.

mutations in CML, combination treatments give patients a better chance of cure than do single-drug treatments. And *in vitro* studies⁷ that compared CML cells treated with one small-molecule inhibitor with those treated with a combination of two or three demonstrated that the combined therapy suppressed cell proliferation more effectively. But with Bozic and colleagues' success in synthesizing theoretical and experimental methods and applying the analysis to solid tumours, these modelling studies have taken a leap forward.

Even more significantly, the authors' paper outlines a roadmap for future personalized therapies, by showing that specific parameters for a patient can be measured and used in a mathematical model to calculate the probability of treatment success and to design the best possible treatment strategy. The authors extracted tumour parameters — including its size at presentation, cell division and cell death rates, and changes in the associated kinetic parameters following treatment — from 20 patients with melanoma who were treated with the small-molecule inhibitor vemurafenib. With this information, they were able to predict the most likely outcome of single, dual and triple therapies for each patient.

There is a bright future for this approach. As new drugs and more information on the exact mechanisms of drug action become available, the model can be iteratively improved. For example, there is currently a strong research focus on cellular plasticity, the heterogeneity of cells within a tumour and the role of

cancer stem cells. But it is not known how the presence of cancer cells with differing properties affects a tumour's susceptibility to targeted treatments. Moreover, the evolutionary costs of resistance for a cell have not been quantified in most cases, nor have mutation rates for molecular changes of various kinds, although estimates have been made for the number of mutations conferring resistance to certain drugs in CML^{7,8}. The potential complications of drug cocktails — including toxicity and undesirable drug interactions — must also be taken into account. But the overall message is loud and clear: mathematics can help to calculate treatment strategies, and the best hope so far lies in combination therapies. ■

Natalia L. Komarova is in the Departments of Mathematics and Ecology and Evolutionary Biology, University of California Irvine, Irvine, California 92697, USA. **C. Richard Boland** is in the Department of Gastroenterology, Baylor University Medical Center, Dallas 75246, Texas, USA.
e-mail: komarova@uci.edu

- Bozic, I. *et al.* *eLIFE* **2**, e00747 (2013).
- Zhang, J., Yang, P. L. & Gray, N. S. *Nature Rev. Cancer* **9**, 28–39 (2009).
- Diaz, L. A. Jr *et al.* *Nature* **486**, 537–540 (2012).
- Katouli, A. A. & Komarova, N. L. *Bull. Math. Biol.* **73**, 549–584 (2011).
- Komarova, N. L. & Wodarz, D. *Proc. Natl Acad. Sci. USA* **102**, 9714–9719 (2005).
- Komarova, N. L., Katouli, A. A. & Wodarz, D. *PLoS ONE* **4**, e4423 (2009).
- Bradeen, H. A. *et al.* *Blood* **108**, 2332–2338 (2006).
- Katouli, A. A. & Komarova, N. L. *PLoS ONE* **5**, e12300 (2010).

THINK BIG

Signs of life.

BY GREGORY BENFORD

When Judy found the strong emission lines of sulphur dioxide in a planetary atmosphere, she knew they had to be wrong. She took the data upstairs.

"A deep sulphur haze covering the whole planet?" her supervisor agreed. "Never happen. Even Krakatoa didn't do that to half an optical depth."

But the measurements held up. The planet was 579 light years away, had a surface temperature of 24°C, and the oxygen-rich atmosphere had deep carbon dioxide lines, too. In shorts and a T-shirt, Judy went for a run along the Charles River to clear her mind. It was 29°C on a clear day, three weeks before Christmas, and the maple trees in Boston hadn't yet turned red.

Back at MIT, she ran a full climate model and found that the planet's star, a G4 type, would heat the surface to 38°C unless she included the sulphur haze. "That's what's keeping the surface at 24°C," she told her professor.

"Um. Any spectral lines from pollution?"

"Sure — carbon dioxide, the sulphur."

"That must be natural somehow."

"Why? The CO₂ keeping Boston warm is from us."

It took another year to prove the case: the planet was being geoengineered with stratospheric sulphur aerosols. "Their civilization went down our same path, burning fossil fuels," Judy said at a press conference at the next meeting of the American Geophysical Union. The press headlined the *Aliens Discovered!* angle, but Judy persisted. "So they're no smarter than we are. Now, they're managing their way out of it."

She didn't like the media circus that erupted and so she fled to Europe. But there, protestors accused her of advocating "alien technology transfer" to Earth. As the world's weather was steadily worsening, this didn't seem to be such a bad idea.

But then the continuing telescopic searches turned up a world that was orbiting beyond the nominal 'habitable zone' of a K star, but that was unaccountably warm. Judy and her

team — by now she had plenty of grants — found CFCs in that atmosphere; the same molecules that had brought grief to Earth's ozone layer in the twentieth century. Careful analysis of the telescopic data revealed a few dozen pixels showing oceans and land

masses with plant covering. Apparently the CFC warming was a full-scale terraforming event, making a cold world habitable.

team — by now she had plenty of grants — found CFCs in that atmosphere; the same molecules that had brought grief to Earth's ozone layer in the twentieth century. Careful analysis of the telescopic data revealed a few dozen pixels showing oceans and land

masses with plant covering. Apparently the CFC warming was a full-scale terraforming event, making a cold world habitable.

Those astronomers keeping track of transits across the disks of stars reported another surprise. They had spotted something bigger than a planet orbiting an F star, but it wasn't round. The oblong profile showed no atmosphere clinging to it, but a shimmer revealed spectral lines implying that the construct was mostly made of carbon.

Building a Dyson sphere? the normally unimaginative *New York Times* headlined. This seemed to be confirmed a few months later when another oblong shell passed across the F star, this time in an orbit farther out than the first, and at a 40° angle to the first orbital plane.

"This shows that aliens are builders," Judy said. But that was not all.

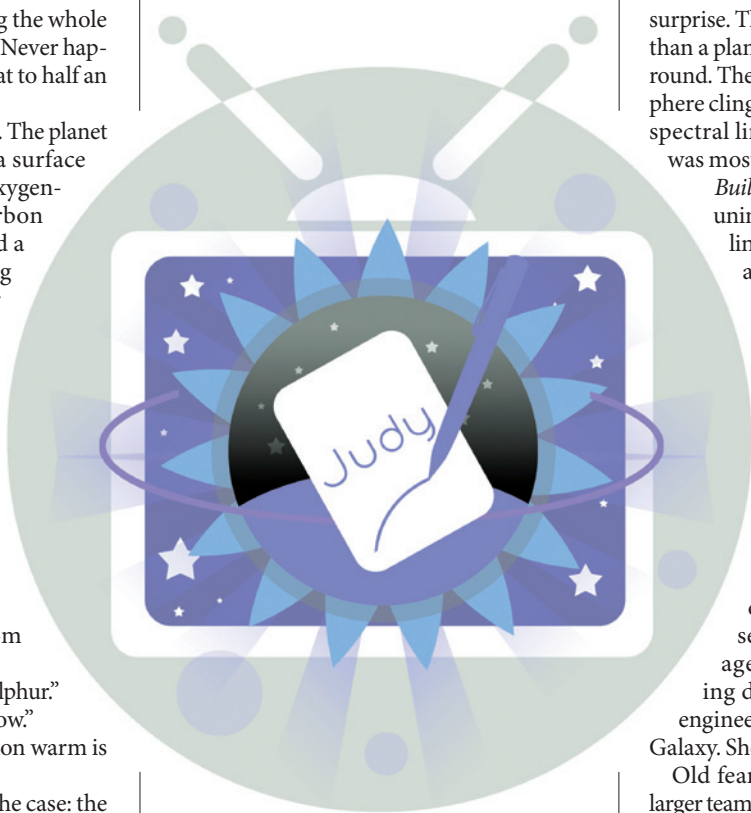
She continued her research, climbing to a top position at the University of California, San Diego. For decades she ran detection strategies for intense searches of billions of stars. She aged well, watching the gathering data, revealing dozens of astro-engineering signatures spread across the Galaxy. She took up surfing.

Old fears resurfaced when Judy's ever-larger team spotted a bright star that suddenly bloomed high above our ecliptic. It was plainly the fusion plume of a decelerating craft. Years later, the slim, magnetically webbed ship went into a high orbit above Earth.

They were spidery aliens who did no climate modification at all. They lived in elaborate cave labyrinths on a cool world that ran on geothermal power, but they liked to rove. They, too, had seen the immense engineering feats in their sky, and so had picked up on Earth's. They were neighbours, from only 23.6 light years away. Because they could directly monitor our low-power TV broadcasts, they knew the entire history of Judy's discoveries. They had become entranced with her.

They had come to talk. And maybe ... an autograph? ■

Gregory Benford is a professor of physics at the University of California, Irvine, and a novelist.



➔ **NATURE.COM**

Follow Futures:

@NatureFutures

go.nature.com/mtoodm

Carbon catabolite repression of the maltose transporter revealed by X-ray crystallography

Shanshuang Chen¹, Michael L. Oldham², Amy L. Davidson³ & Jue Chen^{1,2}

Efficient carbon utilization is critical to the survival of microorganisms in competitive environments. To optimize energy usage, bacteria have developed an integrated control system to preferentially uptake carbohydrates that support rapid growth. The availability of a preferred carbon source, such as glucose, represses the synthesis and activities of proteins necessary for the transport and metabolism of secondary carbon sources. This regulatory phenomenon is defined as carbon catabolite repression¹. In enteric bacteria, the key player of carbon catabolite repression is a component of the glucose-specific phosphotransferase system, enzyme IIA (EIIA^{Glc})^{1,2}. It is known that unphosphorylated EIIA^{Glc} binds to and inhibits a variety of transporters when glucose is available^{1,2}. However, understanding the underlying molecular mechanism has been hindered by the complete absence of structures for any EIIA^{Glc}-transporter complexes. Here we present the 3.9 Å crystal structure of *Escherichia coli* EIIA^{Glc} in complex with the maltose transporter, an ATP-binding cassette (ABC) transporter. The structure shows that two EIIA^{Glc} molecules bind to the cytoplasmic ATPase subunits, stabilizing the transporter in an inward-facing conformation and preventing the structural rearrangements necessary for ATP hydrolysis. We also show that the half-maximal inhibitory concentrations of the full-length EIIA^{Glc} and an amino-terminal truncation mutant differ by 60-fold, consistent with the hypothesis that the amino-terminal region, disordered in the crystal structure, functions as a membrane anchor to increase the effective EIIA^{Glc} concentration at the membrane^{3,4}. Together these data suggest a model of how the central regulatory protein EIIA^{Glc} allosterically inhibits maltose uptake in *E. coli*.

Carbon catabolite repression (CCR) is common in bacteria and higher organisms: up to 5–10% of all bacterial genes are subject to CCR regulation¹. Different organisms use different pathways to achieve CCR. In *E. coli*, CCR is mediated by the modulation of the phosphorylation state of EIIA^{Glc}. EIIA^{Glc} is a member of the phosphoenolpyruvate (PEP)-carbohydrate phosphotransferase system (PTS) that transports and concomitantly phosphorylates several preferred carbohydrates^{1,2}. The influx of PTS substrates and other readily metabolizable carbon sources increases the level of unphosphorylated EIIA^{Glc} by affecting the intracellular [PEP]/[pyruvate] ratio⁵. Phosphorylated EIIA^{Glc} stimulates cAMP synthesis², which leads to transcriptional activation of many catabolic genes. Unphosphorylated EIIA^{Glc} directly inhibits several non-PTS sugar transport systems, including the maltose transporter (MalFGK₂), lactose permease (LacY), melibiose permease (MelB) and raffinose permease (RafB)². It also prevents glycerol uptake by binding to glycerol kinase, an enzyme key to glycerol metabolism². These carbohydrates or their derivatives function as inducers to control the synthesis of the corresponding transporters and metabolic enzymes. Thus, interaction of EIIA^{Glc} with the transporter not only directly prevents secondary carbon source uptake, but also downregulates the expression of the corresponding catabolic systems by reducing the intracellular level of the inducer^{1,2}. The regulation of the *lac* operon is the best known example of this feedback mechanism, known as inducer exclusion^{1,2}. The mechanisms of how EIIA^{Glc} regulates target transporters

have been studied for decades. Mutants of the lactose, maltose, raffinose and melibiose permeases that abolish inducer exclusion have been isolated². The interfaces between EIIA^{Glc} and the target transporters have been probed biochemically². Although those studies have identified some of the residues involved in the interaction, no structure of any transporter in complex with EIIA^{Glc} is yet available.

The maltose transporter MalFGK₂ belongs to the superfamily of ABC transporters. It consists of two transmembrane subunits, MalF and MalG, and two copies of the ATPase subunit, MalK. The structure of MalK can be further divided into an N-terminal nucleotide-binding domain (NBD) and a C-terminal regulatory domain. Crystal structures of the maltose transporter have been determined in three distinct functional states. The inward-facing conformation, also known as the resting state in which the transporter has very low basal ATPase activity, was obtained in the absence of maltose and nucleotides⁶. The outward-facing conformation, corresponding to a hydrolysis-competent intermediate, was stabilized by the periplasmic maltose-binding protein (MBP) and ATP⁷. Finally, a pre-translocation complex, which shows how substrate initiates the transport cycle, was captured in the presence of MBP but in the absence of ATP⁸. These structures illustrate the coordinated motions of the transporter that couple ATP hydrolysis to maltose translocation^{6–8}. On this foundation, the maltose transporter would be a productive system to study how sugar transport is regulated under conditions subject to inducer exclusion.

To obtain crystals of the maltose transporter bound with EIIA^{Glc}, we tested many constructs, including full-length and truncated proteins devoid of flexible regions. It was unclear whether EIIA^{Glc} binds to the transporter in the resting state or an intermediate conformation induced by MBP and/or ATP. Therefore, we carried out crystallization experiments with different combinations of cofactors. Crystals were obtained with full-length MalFGK₂ and EIIA^{Glc} in the absence of MBP and ATP. Initially, the crystals obtained in detergent diffracted X-rays poorly (about 15–20 Å). Crystals grown in lipid/detergent bicelles diffracted to 7 Å resolution. Dehydration caused the unit cell to shrink by 20 Å in one dimension and improved the resolution to 3.9 Å. The structure was determined by molecular replacement and refined with deformable elastic network constraints (Supplementary Table 1).

The asymmetric unit of the crystal contains two copies of the complex, each consisting of one transporter MalFGK₂ and two EIIA^{Glc} molecules (Fig. 1). The maltose transporter shows a conformation similar to that of the resting state⁶. The root mean squared deviation (r.m.s.d.) of the transporter (1,168 C α positions, excluding flexible regions) between the resting state and the EIIA^{Glc} complex structure is 2.2 Å (Supplementary Fig. 1). In this conformation, the transmembrane maltose-binding site is exposed to the intracellular side of the membrane (inward-facing). In contrast to the resting state in which the second periplasmic loop of MalF (P2) is disordered, the structure of the P2 loop is well resolved (Fig. 1 and Supplementary Fig. 2). The relative orientation of the P2 loop is different between the two structures in the asymmetric unit, probably influenced by crystal packing (Supplementary Figs 3 and 4). The two EIIA^{Glc} molecules are bound at

¹Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA. ²Howard Hughes Medical Institute, West Lafayette, Indiana 47907, USA. ³Department of Chemistry, Purdue University, West Lafayette, Indiana 47907, USA.

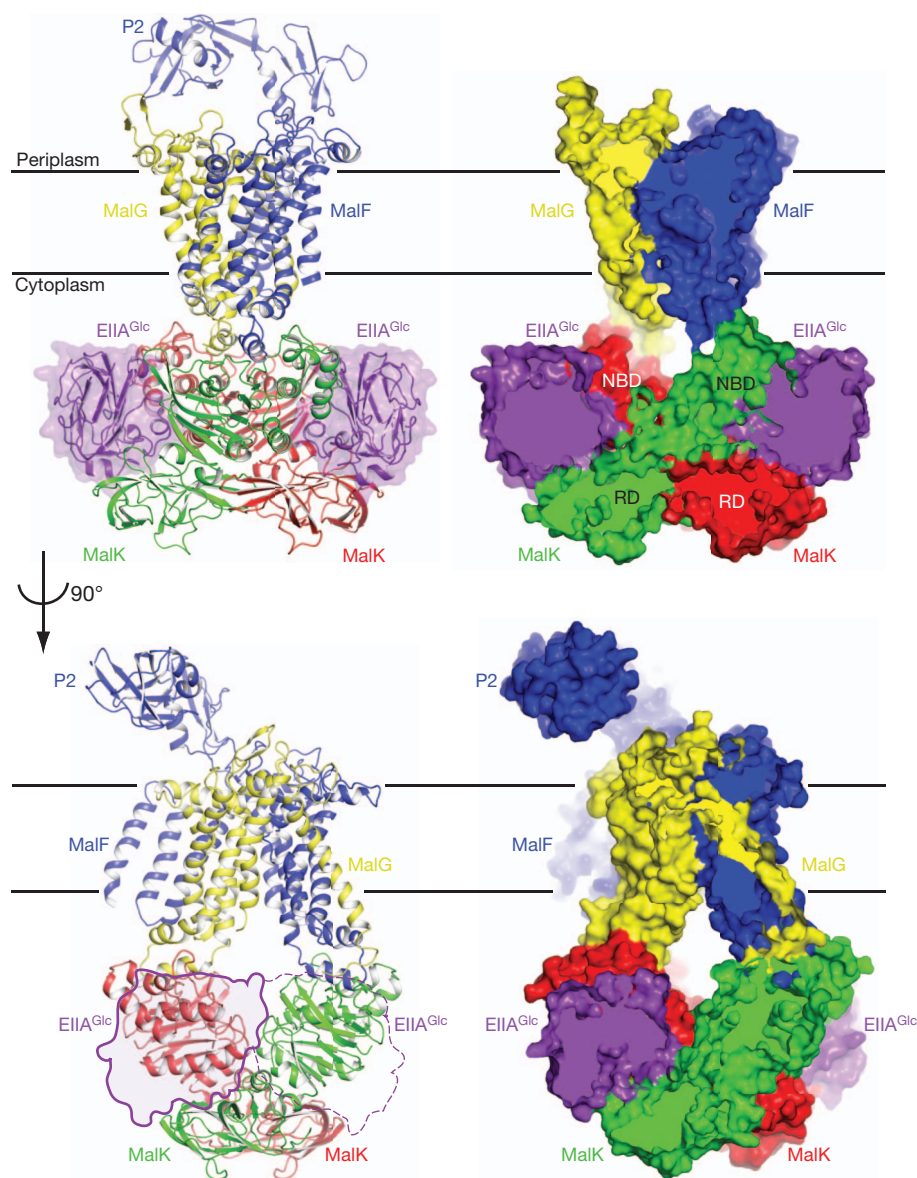


Figure 1 | Two orthogonal views of the EIIGlc-MalFGK₂ complex. Both a ribbon diagram (left) and a slab view (right) are shown. In the lower panel, EIIGlc is represented by a transparent surface, with the solid line representing the front molecule and dotted line representing the backside one. RD, regulatory domain.

structurally equivalent positions, interacting exclusively with the intracellular MalK subunits. The EIIGlc-binding sites are distant from the Walker A motifs that bind ATP (Fig. 2a), the maltose-binding site in MalF and the periplasmic loops to which maltose-MBP docks. Thus, EIIGlc is a classical Monod–Wyman–Changeux allosteric inhibitor of the maltose transporter⁹.

The MalK subunits form an open dimer, interacting with each other through their C-terminal regulatory domains (Fig. 2a). Each EIIGlc is wedged between the NBD of one MalK and the regulatory domain of the opposite MalK, linking these two domains together (Fig. 2a). Of the 1,500 Å² buried surface area on EIIGlc, approximately 55% is buried at the NBD interface and the other 45% at the regulatory domain interface. From this structure, it becomes apparent why binding of EIIGlc would inhibit the maltose transporter. In contrast to the open dimer stabilized by EIIGlc, in the outward-facing state the two NBDs of MalK make contact with each other and two ATPs lie buried along the NBD interface (Fig. 2b). Superposition of the resting state and outward-facing structures shows that the conformational changes of MalK observed in the transport cycle are achieved principally by inter-domain rotations of the NBDs relative to the regulatory domains (Fig. 2b). In effect, the MalK subunits resemble a pair of tweezers, with the regulatory domains acting as the fulcrum point allowing the two

NBDs to open and close like the pincers¹⁰. Here we observe that by linking the NBD of one MalK with the opposite regulatory domain, the two EIIGlc molecules fasten the tweezers in the open configuration. The two transmembrane subunits MalF and MalG are docked onto the MalK NBDs, and conformational changes in the NBDs are tightly coupled to those of the transmembrane domains⁷ (Figs 1 and 2). Thus, by preventing closure of the NBDs, EIIGlc prevents the formation of the outward-facing conformation, a key step for ATP hydrolysis and maltose translocation. In excellent agreement with the crystal structure, all mutations that render resistance to EIIGlc inhibition reside in the MalK subunit^{11,12}. Among the eight mutations isolated, five make direct interactions with EIIGlc in the crystal structure (Fig. 2c). The others, G278P, G302D and S322F, are located in close proximity to the interface and probably produce local conformational changes that prevent EIIGlc from binding (Fig. 2c).

The structure of EIIGlc observed in the complex (residues 19–168) is very similar to that of the free form¹³, with an overall r.m.s.d. of 0.5 Å. This reinforces the observation that neither phosphorylation nor binding to target proteins induces substantial conformational changes of EIIGlc^{14–17}. As a central signalling molecule, EIIGlc interacts with many proteins that share little sequence or structural similarity². The NBD of MalK binds to a canonical surface of EIIGlc that normally

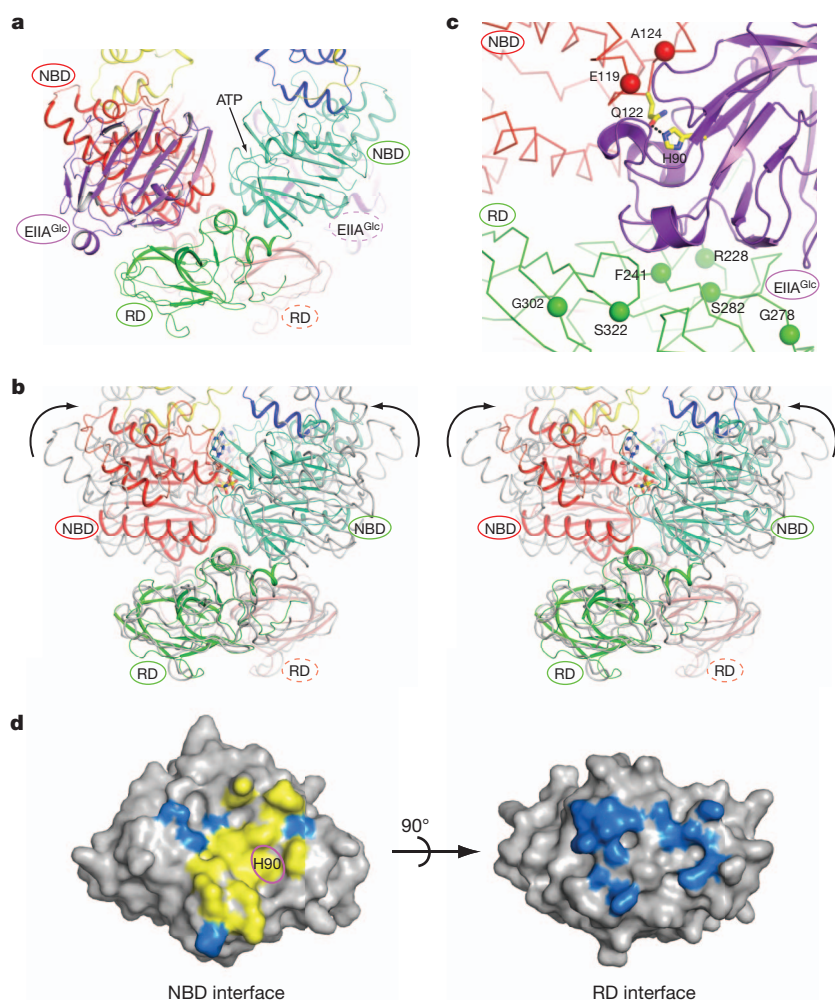


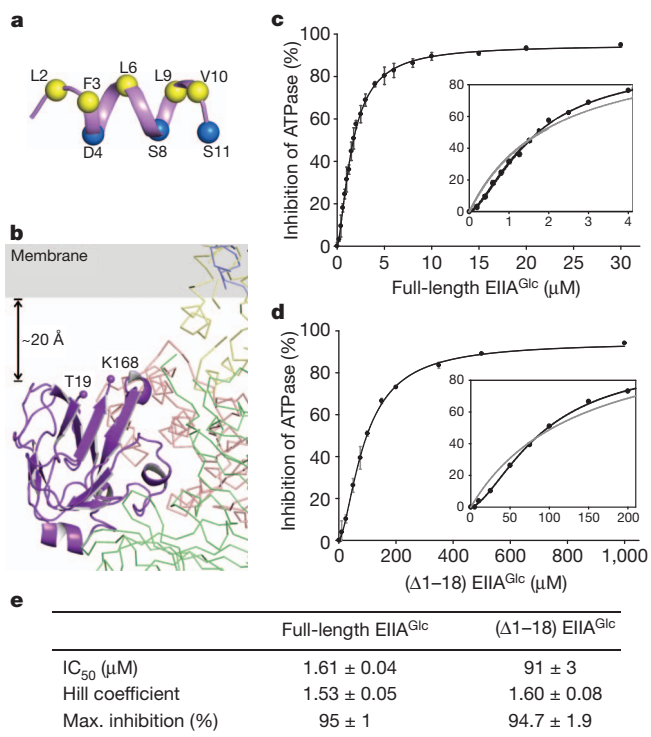
Figure 2 | Binding of EIIA^{Glc} prevents MalK closure. **a**, MalK dimer bound with EIIA^{Glc}. EIIA^{Glc} is shown in purple. The two MalK subunits are coloured in red and green. The NBD and regulatory domain within each MalK subunit are rendered in different shades. The ATP-binding site (the Walker A motif) in the front subunit (green) is indicated by an arrow. **b**, Stereoview of a superposition of the closed MalK dimer, coloured as in panel **a**, with the open dimer coloured in grey. ATP is shown in stick model. The arrows indicate the rotations of the NBDs necessary to form the closed dimer. **c**, EIIA^{Glc}-MalK interface. Inducer-exclusion resistant mutants are labelled. The active site His 90 of EIIA^{Glc} and Gln 122 of MalK are shown in stick models, and the hydrogen bond between His 90 and Gln 122 is indicated by a dashed line. **d**, Surface presentations of EIIA^{Glc}. Residues located within 4.5 Å from the NBD (left) or regulatory domain (right) are coloured. Yellow, the 11 residues (Val 39, Val 40, Phe 41, Ile 45, Val 46, Lys 69, Phe 71, Glu 72, Phe 88, His 90 and Val 96) involved in interacting with all four EIIA^{Glc} partners: MalK, glycerol kinase, HPr and EIIB^{Glc}. The phosphorylation site His 90 is labelled. Blue, MalK-interacting residues that are not part of the canonical binding surface.

interacts with other partners (Fig. 2d, left). Of the 17 residues making contacts with the NBD, 11 are also involved in interacting with glycerol kinase, EIIB^{Glc} and histidine protein (HPr)^{14–16} (Fig. 2d, left). Most of the interface residues are hydrophobic, consistent with the promiscuous interactions this surface makes with different proteins. The crystal structure also shows why phosphorylation prevents the binding of EIIA^{Glc} to the maltose transporter. The phosphorylation site His 90 (ref. 18), located at the EIIA^{Glc}-MalK interface, is in its unphosphorylated form and hydrogen bonded with Gln 122 of MalK (Fig. 2c). Adding a negatively charged phosphate group to the side chain of His 90 would disrupt the interface, thus relieving the maltose transporter from inducer exclusion. As previously suggested by antibody and peptide mapping experiments^{19,20}, residues contacting the regulatory domain are located away from the active site, forming an interface unique to the maltose transporter complex (Fig. 2d, right).

Although the N-terminal 18 residues of EIIA^{Glc} are disordered in the crystal structure, data from other studies suggested that this region might still have an important role in stabilizing the EIIA^{Glc}-MalFGK₂ complex. Studies of synthetic peptides showed that residues 3–10 of EIIA^{Glc} form an amphipathic helix in the presence of either detergent micelles or lipid vesicles (Fig. 3a), suggesting that it may function as a membrane anchor^{3,4}. Furthermore, removing the N-terminal region of EIIA^{Glc} hinders its interaction with membrane transporters, but not soluble partners^{19,21–23}. In the EIIA^{Glc}-MalFGK₂ structure, EIIA^{Glc} is oriented such that the N terminus is pointed towards the membrane. The first ordered residue Thr 19 is approximately 20 Å outside the lipid bilayer (Fig. 3b). To test the function of the N-terminal region, we

compared the inhibition of maltose-MBP stimulated ATPase activity by full-length EIIA^{Glc} and a truncation mutant consisting of residues 19–168 (Δ 1–18). The maltose transporter was reconstituted into lipid nanodiscs to allow access of MBP and EIIA^{Glc} to both sides of the membrane. Both constructs of EIIA^{Glc} inhibit with positive cooperativity (Fig. 3c, d), consistent with the 2:1 stoichiometry of the complex and an earlier observation that inhibition of maltose uptake was sigmoidally related to EIIA^{Glc} concentration²⁴. However, the half-maximal inhibitory concentration (IC₅₀) of the Δ 1–18 construct is 60 times higher than that of the full-length protein (91 μ M versus 1.6 μ M, Fig. 3e). This finding is consistent with the hypothesis that the full-length protein is concentrated near the membrane through the N-terminal region. Given that the intracellular concentration of EIIA^{Glc} is around 50 μ M²⁵, the membrane-anchoring function of the N-terminal region would be very important to ensure efficient inhibition of the transporter. The high IC₅₀ value of the Δ 1–18 construct also explains the lack of inhibition observed with a similar construct from a previous study, where the concentration used in the assay was approximately 24 μ M¹⁹.

On the basis of the structural and functional data, we propose a model for inducer exclusion of the maltose transporter (Fig. 4). The uptake and concomitant phosphorylation of glucose lead to dephosphorylation of EIIA^{Glc} through a cascade of enzymes. Unphosphorylated EIIA^{Glc} binds to the resting state maltose transporter and locks the MalK dimer in the open configuration. Ordinarily, binding of the maltose-loaded MBP and ATP are sufficient to promote the transition from the inward-facing, resting state to the outward-facing, ATPase-active state, the very conformational change that allows maltose entry from the periplasm



and also enables ATP hydrolysis at the same time. Under inducer exclusion, EIIA^{Glc} stabilizes the resting state such that MBP and ATP binding can no longer initiate the transport cycle. The maltose

Figure 3 | The N terminus of EIIA^{Glc} probably functions as a membrane anchor. **a**, Structure of the N-terminal peptide determined by NMR in dihexanoyl phosphatidylglycerol (PDB 1O53)⁴. Hydrophobic residues lining one side of the helix are shown in yellow. Hydrophilic residues on the opposite side of the helix are shown in blue. **b**, Close-up view of EIIA^{Glc} bound to MalK. The two terminal residues visible in the crystal structure (Thr 19 and Lys 168) are labelled. **c, d**, Inhibition of the MBP-stimulated ATPase activity of reconstituted nanodiscs by full-length (**c**) and N-terminally truncated EIIA^{Glc} (**d**). Data points represent the means \pm standard deviation of triplicate measurements. Insets show enlargements of the figures fitted to the Hill equation (black line) and to the Michaelis–Menten equation (grey line). **e**, Summary table of the kinetics.

transporter is the only ABC transporter known to be inhibited by EIIA^{Glc} ; however, similar allosteric regulation mechanisms by the transported substrate itself have been described for the methionine transporter and the molybdate transporter^{26,27}. In both cases, at high intracellular concentrations, methionine or molybdate binds to allosteric sites in the ATPase subunits to stabilize the transporter in the inward-facing state, thus preventing further substrate uptake^{26,27}.

Previously it was shown that binding of EIIA^{Glc} to two other transporters, LacY and MelB, were enhanced by their corresponding sugar substrates^{28,29}. However, for the maltose transporter no cooperativity between substrate and EIIA^{Glc} binding was observed³⁰. Consistent with the functional data, the crystal structure offers no evidence that the EIIA^{Glc}-MalFGK₂ interaction would be affected by the presence of maltose. LacY and MelB are homologous proteins belonging to the major facilitator superfamily, unrelated to ABC transporters. Thus, it is probable that the mechanism underlying LacY and MelB regulation is different from that of the maltose transporter.

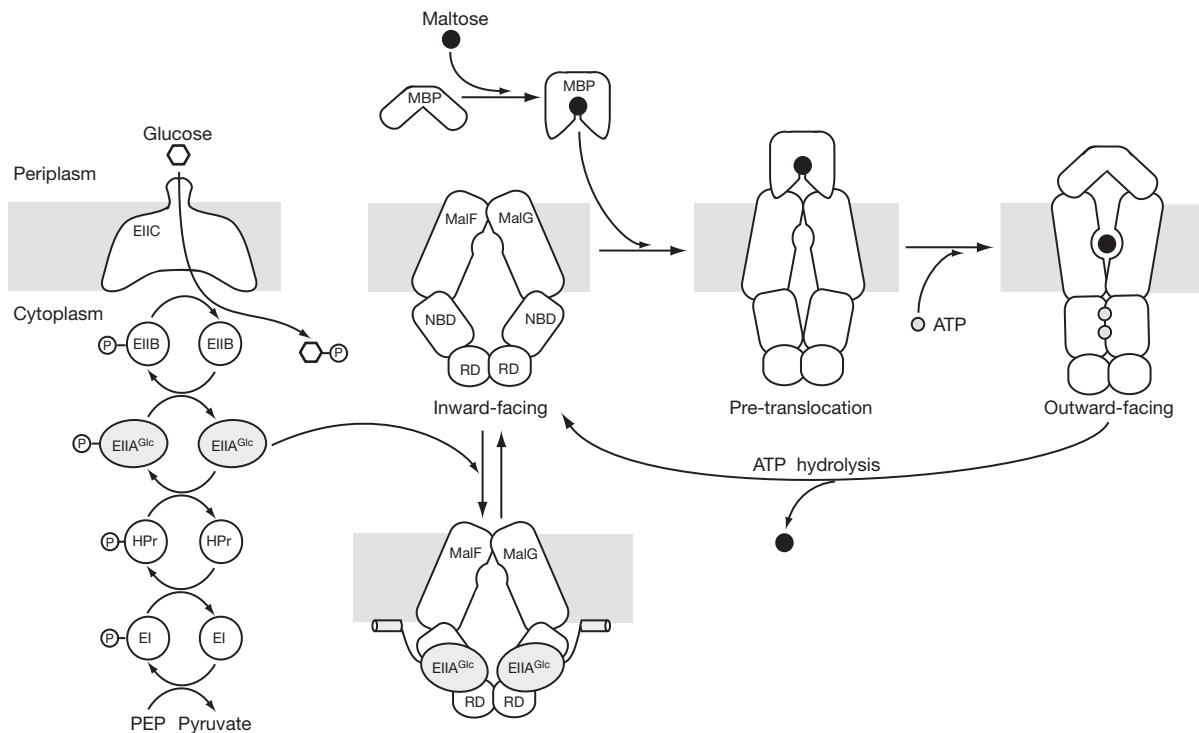


Figure 4 | Inhibition of the maltose transporter by inducer exclusion. Left, glucose and other PTS substrates are transported and phosphorylated by five coupled reactions, leading to an increased level of unphosphorylated EIIA^{Glc}. Right, the maltose transport cycle. In the resting state, the open MalK dimer coincides with the inward-facing transmembrane domains. Upon association of MBP, the transporter undergoes a conformational change (the pre-translocation state) that permits ATP to promote a concerted motion of MalK

closure, reorientation of the transmembrane domains and opening of MBP. Formation of the outward-facing conformation transfers maltose from MBP to the transmembrane binding site, and at the same time positions ATP at the catalytic site for hydrolysis. ATP hydrolysis releases maltose into the cytoplasm and resets the transporter to the resting state. Under conditions subject to inducer exclusion, unphosphorylated EIIA^{Glc} binds to and stabilizes the resting state transporter, thus inhibiting maltose transport.

METHODS SUMMARY

EIIA^{Glc} and MalFGK₂ were expressed separately in *E. coli*. MalFGK₂ was purified by metal-affinity and gel-filtration chromatography using *n*-dodecyl- β -D-maltoside. EIIA^{Glc} was purified using metal affinity, ion-exchange and gel-filtration chromatography. Protein sample composed of EIIA^{Glc} and MalFGK₂ mixed at 2.5/1 molar ratio was incubated with 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine/3-(3-cholamidopropyl)dimethylammonio-2-hydroxy-1-propanesulphonate (DMPC/CHAPSO) bicelles before crystallization at 22 °C. Diffraction data were collected at the Advanced Photon Source. The structure was solved using molecular replacement (Phaser, CCP4) using structures of isolated EIIA^{Glc} (PDB 1F3G) and the resting state MalFGK₂ (PDB 3FH6). The model was built in Coot, and refined in CNS and PHENIX. The ATPase activity was measured by recording the absorption at 340 nm in an ATP regeneration-NADH consumption-coupled system, using reconstituted MalFGK₂ in nanodiscs. Triplicated data were fit to either the Hill or the Michaelis-Menten equation using SigmaPlot (SYSTAT Software).

Full Methods and any associated references are available in the online version of the paper.

Received 5 February; accepted 29 April 2013.

Published online 16 June 2013.

- Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature Rev. Microbiol.* **6**, 613–624 (2008).
- Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol. Mol. Biol. Rev.* **70**, 939–1031 (2006).
- Wang, G., Peterkofsky, A. & Clow, G. M. A novel membrane anchor function for the N-terminal amphipathic sequence of the signal-transducing protein IIA^{Glucose} of the *Escherichia coli* phosphotransferase system. *J. Biol. Chem.* **275**, 39811–39814 (2000).
- Wang, G., Keifer, P. A. & Peterkofsky, A. Solution structure of the N-terminal amphitropic domain of *Escherichia coli* glucose-specific enzyme IIA in membrane-mimetic micelles. *Protein Sci.* **12**, 1087–1096 (2003).
- Hogema, B. M. *et al.* Inducer exclusion in *Escherichia coli* by non-PTS substrates: the role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIA^{Glc}. *Mol. Microbiol.* **30**, 487–498 (1998).
- Khare, D., Oldham, M. L., Orelle, C., Davidson, A. L. & Chen, J. Alternating access in maltose transporter mediated by rigid-body rotations. *Mol. Cell* **33**, 528–536 (2009).
- Oldham, M. L., Khare, D., Quijcho, F. A., Davidson, A. L. & Chen, J. Crystal structure of a catalytic intermediate of the maltose transporter. *Nature* **450**, 515–521 (2007).
- Oldham, M. L. & Chen, J. Crystal structure of the maltose transporter in a pretranslocation intermediate state. *Science* **332**, 1202–1205 (2011).
- Monod, J., Wyman, J. & Changeux, J. P. On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* **12**, 88–118 (1965).
- Chen, J., Lu, G., Lin, J., Davidson, A. L. & Quijcho, F. A. A tweezers-like motion of the ATP-binding cassette dimer in an ABC transport cycle. *Mol. Cell* **12**, 651–661 (2003).
- Dean, D. A., Reizer, J., Nikaïdo, H. & Saier, M. H. Jr. Regulation of the maltose transport system of *Escherichia coli* by the glucose-specific enzyme III of the phosphoenolpyruvate-sugar phosphotransferase system. Characterization of inducer exclusion-resistant mutants and reconstitution of inducer exclusion in proteoliposomes. *J. Biol. Chem.* **265**, 21005–21010 (1990).
- Kühnau, S., Reyes, M., Sievertsen, A., Shuman, H. A. & Boos, W. The activities of the *Escherichia coli* MalK protein in maltose transport, regulation, and inducer exclusion can be separated by mutations. *J. Bacteriol.* **173**, 2180–2186 (1991).
- Worthylake, D. *et al.* Three-dimensional structure of the *Escherichia coli* phosphocarrier protein IIGlc. *Proc. Natl Acad. Sci. USA* **88**, 10382–10386 (1991).
- Cai, M. *et al.* Solution structure of the phosphoryl transfer complex between the signal-transducing protein IIA^{Glucose} and the cytoplasmic domain of the glucose transporter IICB^{Glucose} of the *Escherichia coli* glucose phosphotransferase system. *J. Biol. Chem.* **278**, 25191–25206 (2003).
- Hurley, J. H. *et al.* Structure of the regulatory complex of *Escherichia coli* IIGlc with glycerol kinase. *Science* **259**, 673–677 (1993).
- Wang, G. *et al.* Solution structure of the phosphoryl transfer complex between the signal transducing proteins HPr and IIA^{Glucose} of the *Escherichia coli* phosphoenolpyruvate:sugar phosphotransferase system. *EMBO J.* **19**, 5635–5649 (2000).
- Pelton, J. G., Torchia, D. A., Meadow, N. D. & Roseman, S. Structural comparison of phosphorylated and unphosphorylated forms of IIGlc, a signal-transducing protein from *Escherichia coli*, using three-dimensional NMR techniques. *Biochemistry* **31**, 5215–5224 (1992).
- Dörschug, M., Frank, R., Kalbitzer, H. R., Hengstenberg, W. & Deutscher, J. Phosphoenolpyruvate-dependent phosphorylation site in enzyme IIGlc of the *Escherichia coli* phosphotransferase system. *Eur. J. Biochem.* **144**, 113–119 (1984).
- Blischke, B., Volkmer-Engert, R. & Schneider, E. Topography of the surface of the signal-transducing protein EIIA^{Glc} that interacts with the MalK subunits of the maltose ATP-binding cassette transporter (MalFGK₂) of *Salmonella typhimurium*. *J. Biol. Chem.* **281**, 12833–12840 (2006).
- Stein, A. *et al.* Functional characterization of the maltose ATP-binding-cassette transporter of *Salmonella typhimurium* by means of monoclonal antibodies directed against the MalK subunit. *Eur. J. Biochem.* **269**, 4074–4085 (2002).
- Meadow, N. D. & Roseman, S. Sugar transport by the bacterial phosphotransferase system. Isolation and characterization of a glucose-specific phosphocarrier protein (IIGlc) from *Salmonella typhimurium*. *J. Biol. Chem.* **257**, 14526–14537 (1982).
- Meadow, N. D., Savtchenko, R. S., Remington, S. J. & Roseman, S. Effects of mutations and truncations on the kinetic behavior of IIA^{Glc}, a phosphocarrier and regulatory protein of the phosphoenolpyruvate phosphotransferase system of *Escherichia coli*. *J. Biol. Chem.* **281**, 11450–11455 (2006).
- Misko, T. P., Mitchell, W. J., Meadow, N. D. & Roseman, S. Sugar transport by the bacterial phosphotransferase system. Reconstitution of inducer exclusion in *Salmonella typhimurium* membrane vesicles. *J. Biol. Chem.* **262**, 16261–16266 (1987).
- van der Vlag, J., van Dam, K. & Postma, P. W. Quantification of the regulation of glycerol and maltose metabolism by IAGlc of the phosphoenolpyruvate-dependent glucose phosphotransferase system in *Salmonella typhimurium*. *J. Bacteriol.* **176**, 3518–3526 (1994).
- Scholte, B. J., Schuitema, A. R. & Postma, P. W. Isolation of IIGlc of the phosphoenolpyruvate-dependent glucose phosphotransferase system of *Salmonella typhimurium*. *J. Bacteriol.* **148**, 257–264 (1981).
- Kadaba, N. S., Kaiser, J. T., Johnson, E., Lee, A. & Rees, D. C. The high-affinity *E. coli* methionine ABC transporter: structure and allosteric regulation. *Science* **321**, 250–253 (2008).
- Gerber, S., Comellas-Bigler, M., Goetz, B. A. & Locher, K. P. Structural basis of trans-inhibition in a molybdate/tungstate ABC transporter. *Science* **321**, 246–250 (2008).
- Osumi, T. & Saier, M. H. Jr. Mechanism of regulation of the lactose permease by the phosphotransferase system in *Escherichia coli*: evidence for protein-protein interaction. *Ann. Microbiol.* **133**, 269–273 (1982).
- Osumi, T. & Saier, M. H. Jr. Regulation of lactose permease activity by the phosphoenolpyruvate:sugar phosphotransferase system: evidence for direct binding of the glucose-specific enzyme III to the lactose permease. *Proc. Natl Acad. Sci. USA* **79**, 1457–1461 (1982).
- Saier, M. H. Jr., Novotny, M. J., Comeau-Fuhrman, D., Osumi, T. & Desai, J. D. Cooperative binding of the sugar substrates and allosteric regulatory protein (enzyme IIGlc of the phosphotransferase system) to the lactose and melibiose permeases in *Escherichia coli* and *Salmonella typhimurium*. *J. Bacteriol.* **155**, 1351–1357 (1983).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the staff at the Advance Photon Source GM/CA-CAT, NE-CAT and SBC for assistance with data collection. This work was supported by a National Institutes of Health grant (GM070515 to J.C. and A.L.D.).

Author Contributions All authors designed the study and analysed the data. S.C. crystallized the complex and performed the biochemical experiments. S.C. and M.L.O. determined the crystal structure and made the figures. J.C. wrote the manuscript with inputs from all authors.

Author Information Coordinates and structure factors have been deposited in the Protein Data Bank under accession number 4JBW. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.C. (chenjue@purdue.edu).

METHODS

Expression and purification of MalFGK₂. Expression and purification of the maltose transporter MalFGK₂ with a 6×-His tag at the C terminus of MalK were carried out as described previously⁷. In brief, the HN741 host strain (*E. coli* K-12 *argH his tpsL1 malT(Con) malBA13 ΔatpBC ilv::Tn10/F' lacIqTn5*) containing chromosomal deletions of the genes of the maltose transport system was transformed simultaneously with three compatible antibiotic-resistance plasmids carrying *malF/malG* (Ampicillin^R), *malK* (Chloramphenicol^R) and *lacIq* (Spectinomycin^R). Cells were grown in Terrific Broth (Novagen) supplemented with ampicillin, chloramphenicol and spectinomycin (Sigma) at 37 °C until $D_{600\text{ nm}} = 0.7\text{--}0.9$ and then induced with 50 μM isopropyl-β-D-thiogalactoside (IPTG) at 22 °C for 20 h before collecting by centrifugation at 4,000g for 15 min. Re-suspended cells were broken by two passes through a high-pressure homogenizer (Emulsiflex-C3; Avestin) and centrifuged at 80,000g for 40 min to isolate the *E. coli* membrane fraction. After re-suspending the membranes in buffer containing 20 mM Tris-HCl, pH 8.0, 100 mM NaCl, 5 mM MgCl₂, 10% glycerol and 0.3% *n*-dodecyl-β-D-maltopyranoside (DDM, Anatrace) at a total protein concentration of 5 mg ml⁻¹, the extracted maltose transporter was purified by cobalt-affinity (Clontech) and gel-filtration (Superdex 200, GE Healthcare) chromatography in buffer containing 10 mM Tris-HCl, pH 8.0, 200 mM NaCl and 0.012% DDM.

Expression and purification of EIIA^{Glc}. Genes encoding the full-length or (Δ1–18) EIIA^{Glc} were cloned into a pMCSG7 LIC vector with a 10×-His tag followed by a tobacco etch virus (TEV) protease cleavage site at the N terminus of the protein. The *E. coli* strain BL21 (DE3) Star was transformed and grown in Luria-Bertani medium (MP Biomedicals) at 37 °C until $D_{600\text{ nm}} = 0.6\text{--}0.8$ before induction by 100 μM IPTG at 16 °C for 20 h. EIIA^{Glc} was purified by cobalt-affinity chromatography in buffer containing 20 mM Tris-HCl, pH 8.0, and 150 mM NaCl, and then incubated with TEV protease to remove the N-terminal 10×-His tag. Tag-free EIIA^{Glc} was further purified by ion-exchange (Source 15Q, GE Healthcare) and gel-filtration (Superdex 75, GE Healthcare) chromatography in buffer containing 10 mM Tris-HCl, pH 8.0, 200 mM NaCl.

Expression and purification of MBP. Expression and purification of wild-type MBP was carried out similarly to that of EIIA^{Glc} by metal-affinity and gel-filtration chromatography.

Crystallization of EIIA^{Glc}–MalFGK₂ complex. Preparation of a 35% bicelle stock was carried out as described previously^{31,32}. In brief, 0.26 g 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC, Avanti Polar Lipids) and 0.09 g 3-(cholamidopropyl)-dimethylammonio-2-hydroxyl-1-propanesulphonate (CHAPSO, Affymetrix) were dissolved in nanopure water at a final volume of 1 ml. Multiple cycles of incubation at 42 °C, vortexing, incubation on ice, and vortexing were performed to obtain a homogeneous bicelle stock that exhibited a temperature-induced phase transition between a transparent viscous gel at 22 °C and a milky solution on ice. A pre-chilled protein sample containing purified MalFGK₂ and EIIA^{Glc} at 1/2.5 molar ratio was mixed with the bicelle stock at a 4/1 ratio (v/v) on ice. The final sample for crystallization was composed of 10 mg ml⁻¹ MalFGK₂, 2.5 mg ml⁻¹ EIIA^{Glc}, 0.01% DDM and 7% bicelles.

Crystals were grown at 22 °C by vapour diffusion using reservoir solution containing 100 mM sodium cacodylate (Sigma), pH 5.6, 13% PEG monomethylether 2000 (Sigma) and 100 mM NDSB-256 (non-detergent sulfobetaine; Hampton Research). Dehydration was performed by replacing the reservoir solution every 24 h with increasing PEG (5% per step) to a final concentration of 35%. After dehydration, crystals were looped and flash frozen in liquid nitrogen.

Data collection and structure determination. X-ray diffraction data were collected at the Advanced Photon Source (APS-23ID) at 100 K. Diffraction images were processed and scaled with HKL2000 (HKL Research). The structure of EIIA^{Glc}–MalFGK₂ complex was solved via molecular replacement³³ (Phaser, CCP4 (ref. 34)) using structures of isolated EIIA^{Glc} (PDB: 1F3G)¹³ and the resting state MalFGK₂

(PDB: 3FH6)⁶. The initial model was refined with deformable elastic network (DEN) constraints³⁵ in CNS³⁶, and then improved by cycles of manual building in Coot³⁷ and refinement in PHENIX³⁸.

Reconstitution of MalFGK₂ into nanodiscs. Reconstitution of MalFGK₂ into nanodiscs was performed as described previously³⁹. In brief, a lipid stock (50 mg ml⁻¹) was prepared from chloroform-dissolved *E. coli* total lipids (Avanti Polar Lipids) after drying under Argon and re-suspension into buffer containing 50 mM Tris-HCl, pH 7.5. Dissolved lipids were purged with Argon and then sonicated to clarity. The membrane scaffold protein (MSP) was expressed and purified as described³⁹. To prepare the nanodiscs, a mixture of *E. coli* lipids, sodium cholate, purified MalFGK₂ and MSP at 120 μg ml⁻¹, 25 mM, 10 μM and 50 μM concentrations, respectively, was incubated at 22 °C for 1 h with gentle rocking. Nanodiscs were assembled using a 5/1 molar ratio of MSP/MalFGK₂ to minimize the incorporation of more than one MalFGK₂ complex per nanodisc. Nanodisc self-assembly was then initiated by adding hydrated Bio-Beads SM-2 (Bio-Rad) to remove the detergent. Bio-Beads were then removed after 3 h incubation at 22 °C. The mixture was loaded onto cobalt affinity resin pre-equilibrated with 20 mM Tris-HCl, pH 8.0, and 100 mM NaCl. The resin was washed extensively to remove MalFGK₂-free nanodiscs. Nanodiscs containing MalFGK₂ were eluted and loaded on a desalting column (NAP-10 Sephadex G-25 columns, GE Healthcare) to remove imidazole.

Kinetic characterization of inhibition of the maltose transporter ATPase by EIIA^{Glc}. The ATPase activity was determined using an ATP regeneration–NADH consumption-coupled system^{40,41} and reconstituted MalFGK₂ in nanodiscs. Different amounts of EIIA^{Glc} were added into 100 μl reaction mixture containing 50 mM HEPES, pH 8.0, 94 nM MalFGK₂, 15 μM MBP, 1 mM maltose, 60 μg ml⁻¹ pyruvate kinase, 32 μg ml⁻¹ lactate dehydrogenase, 4 mM PEP, 0.3 mM NADH and 10 mM MgCl₂. After incubation for 2 h at 25 °C, 1 mM ATP was added to initiate the reaction. The absorbance at 340 nm was recorded using the 8453 UV-Vis Diode Array System (Agilent Technologies) at 25 °C for 5 min. All measurements were performed three times from the same protein preparations. Data were fit to either the Hill or the Michaelis–Menten equation using SigmaPlot (SYSTAT Software).

Figure preparation. All structure figures were prepared with the program PyMOL (<http://www.pymol.org>).

- Ujwal, S. & Bowie, J. U. Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *J. Mol. Biol.* **316**, 1–6 (2002).
- Ujwal, R. & Abramson, J. High-throughput crystallization of membrane proteins using the lipidic bicelle method. *J. Vis. Exp.* **59**, e3383 (2012).
- Rossmann, M. G. The molecular replacement method. *Acta Crystallogr. A* **46**, 73–82 (1990).
- Collaborative Computational Project, number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Schröder, G. F., Levitt, M. & Brunger, A. T. Super-resolution biomolecular crystallography with low-resolution data. *Nature* **464**, 1218–1222 (2010).
- Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Alvarez, F. J., Orelle, C. & Davidson, A. L. Functional reconstitution of an ABC transporter in nanodiscs for use in electron paramagnetic resonance spectroscopy. *J. Am. Chem. Soc.* **132**, 9513–9515 (2010).
- Schachtschmidt, B. F., Keefe, E. B., Blankenship, N. M. & Ockner, R. K. Validation of a recording spectrophotometric method for measurement of membrane-associated Mg- and NaK-ATPase activity. *J. Lab. Clin. Med.* **93**, 790–799 (1979).
- Orelle, C., Ayvaz, T., Everly, R. M., Klug, C. S. & Davidson, A. L. Both maltose-binding protein and ATP are required for nucleotide-binding domain closure in the intact maltose ABC transporter. *Proc. Natl Acad. Sci. USA* **105**, 12837–12842 (2008).

Reversal of an ancient sex chromosome to an autosome in *Drosophila*

Beatriz Vicoso¹ & Doris Bachtrog¹

Although transitions of sex-determination mechanisms are frequent in species with homomorphic sex chromosomes^{1–3}, heteromorphic sex chromosomes are thought to represent a terminal evolutionary stage owing to chromosome-specific adaptations such as dosage compensation or an accumulation of sex-specific mutations^{1,4}. Here we show that an autosome of *Drosophila*, the dot chromosome, was ancestrally a differentiated X chromosome. We analyse the whole genome of true fruitflies (Tephritidae), flesh flies (Sarcophagidae) and soldier flies (Stratiomyidae) to show that genes located on the dot chromosome of *Drosophila* are X-linked in outgroup species, whereas *Drosophila* X-linked genes are autosomal. We date this chromosomal transition to early drosophilid evolution by sequencing the genome of other Drosophilidae. Our results reveal several puzzling aspects of *Drosophila* dot chromosome biology to be possible remnants of its former life as a sex chromosome, such as its minor feminizing role in sex determination⁵ or its targeting by a chromosome-specific regulatory mechanism⁶. We also show that patterns of biased gene expression of the dot chromosome during early embryogenesis, oogenesis and spermatogenesis resemble that of the current X chromosome. Thus, although sex chromosomes are not necessarily evolutionary end points and can revert back to an autosomal inheritance, the highly specialized genome architecture of this former X chromosome suggests that severe fitness costs must be overcome for such a turnover to occur.

Sex is an important and conserved feature, yet sex-determination mechanisms are labile in many taxa with non-differentiated, homomorphic sex chromosomes, and often vary among closely related species or among individuals within a species¹. Highly differentiated, heteromorphic sex chromosomes (that is, a degenerate, gene-poor Y chromosome and an often dosage-compensated X), however, seem to represent an evolutionary end point and become a permanent fixture of the genome in many species groups^{1,3,4}. Special adaptations on highly evolved sex chromosomes—such as dosage compensation or inactivation during male meiosis, an accumulation of sex-determining and sex-specific genes, or inviability of YY individuals—prevent the reversal of heteromorphic sex chromosomes back to autosomes^{1,4}, and sex-chromosome differentiation is viewed as an evolutionary one-way street. Indeed, the heteromorphic sex chromosomes of both mammals and birds originated independently from an ancestor with homomorphic sex chromosomes over 100 million years ago, and the stable inheritance of the sex chromosomes in these two clades reflects their highly specialized genome architecture⁷. By contrast, homomorphic sex chromosomes in fish, amphibians and reptiles often show high rates of turnover between species^{2,3}. Although observations in vertebrates support the notion that heteromorphic sex chromosomes present an ‘evolutionary trap’ inert to turnover⁷, little is known about such transitions in other taxa. Here, we uncover a sex-chromosome reversion within Diptera (flies) in the genetic model organism *Drosophila*.

In many higher Diptera (suborder Brachycera), including *Drosophila*, the basic karyotype ($2n = 12$) consists of 5 pairs of large euchromatic rods (named Muller elements A–E in *Drosophila*; each containing well over 2,000 genes) and a smaller heterochromatic dot chromosome⁸

(element F in *Drosophila*; containing only about 100 genes; see Fig. 1). The gene content of Muller elements is highly conserved across Diptera⁹. One of the large rods (element A) segregates as the X chromosome in *Drosophila*, and sex is determined by the dose of the X-linked gene *Sex lethal* (*Sxl*), with diploid XX embryos developing into females and haploid XY into males¹⁰. The identity of the sex chromosomes and the master sex-determination gene is not conserved across other Diptera families¹¹; instead, a small and often heterochromatic chromosome pair segregates as the sex chromosomes in several other Brachycera¹². No genes have yet been isolated from the sex chromosomes in outgroup species, and the relationship with *Drosophila* chromosomes, if any, is unclear.

We used whole-genome sequencing to identify sex-linked genes in species from several higher Diptera families, including the black soldier fly *Hermetia illucens* (Stratiomyidae, a basal Brachycera), the olive fruitfly *Bactrocera oleae* (Tephritidae), the grey fleshfly *Sarcophaga bullata* (Sarcophagidae), the zoophilic fruitfly *Phortica variegata* (Steganinae, a sister clade to *Drosophila* within Drosophilidae), as well as the basal *Drosophila* species *Drosophila busckii* (Fig. 1 and Supplementary Fig. 1). Paired-end Illumina genomic reads were obtained from single males and females of each species, assembled, and the resulting scaffolds were mapped to the different Muller elements according to

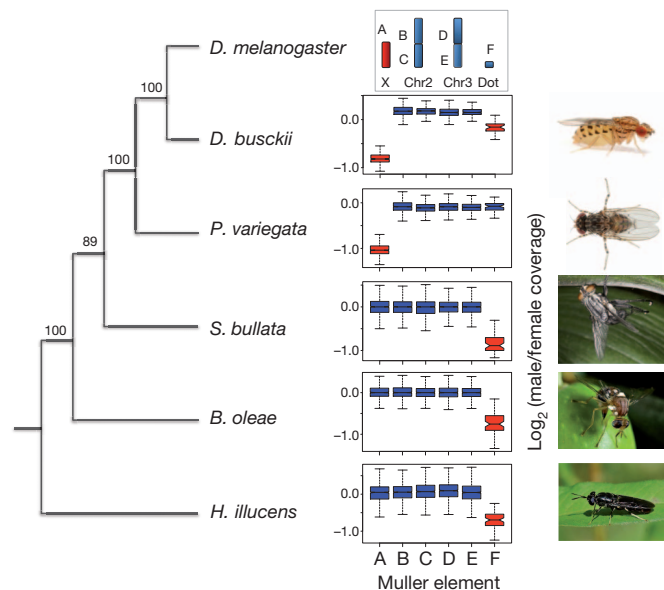


Figure 1 | Sex chromosomes in higher Diptera revealed by genome analysis. Evolutionary relationship inferred from 185 conserved protein-coding genes (93,134 amino acids) using PhyML (with bootstrap values indicated at the nodes), and male-to-female coverage ratio across chromosome elements (Muller elements A–F) in the Diptera species studied. X chromosomes (red) have only half the read coverage in males versus females. Boxes extend from the first to the third quartile and whiskers to the most extreme data point within 1.5 times the interquartile range.

¹Department of Integrative Biology, Center for Theoretical Evolutionary Genomics, University of California Berkeley, Berkeley, California 94720, USA.

their homology to *Drosophila melanogaster* genes (Supplementary Table 1). Male and female reads were mapped back to the scaffolds, and male-to-female coverage was used to detect sex-linked sequences, with X-linked scaffolds only showing half the coverage in males compared to females¹³. Consistent with Muller-A being the ancestral sex chromosome in all Drosophilidae, male-to-female coverage is only half that relative to autosomes in *D. busckii* and *P. variegata* (Fig. 1, Supplementary Table 2 and Supplementary Figs 2–4). In the more distantly related outgroups *B. oleae*, *S. bullata* and *H. illucens*, male-to-female coverage shows that element A is an autosome; instead, element F, the ‘dot chromosome’ of *Drosophila*, has significantly reduced male-to-female coverage in outgroup species (Fig. 1, Supplementary Table 2 and Supplementary Figs 5–7). This indicates that element F is a heteromorphic sex-chromosome pair in several Brachycera, with a completely degenerate Y. In fact, the phylogenetic pattern suggests that the dot was a sex chromosome in an ancestor of *Drosophila*, and only reverted to an autosome in the lineage leading to Drosophilidae.

Certain features of their genome architecture distinguish X chromosomes from autosomes: (1) they contain sex-determining factors (*Sxl* in *Drosophila*)¹⁰; (2) they show a non-random gene content (a deficiency of male- and excess of female-biased genes in *Drosophila*)¹⁴; (3) they are transcriptionally inactivated during spermatogenesis¹⁵; (4) chromosome haploidy is not lethal; and (5) chromosome-wide regulatory mechanisms ensure balanced levels of gene expression in the heterogametic sex (the male-specific lethal (MSL) complex in *Drosophila*)¹⁶. Interestingly, the dot chromosome of *D. melanogaster* harbours several peculiar characteristics that have puzzled drosophilists for decades, but can now be understood as possible remnants of its history as an X chromosome. Specifically, the dot has a minor role in sex determination and contains feminizing factors⁵, with increased dosage of the dot chromosome shifting 2X:3A (A denotes autosome) intersex individuals towards female development⁵. Consistent with its feminizing effect, genes located on the dot chromosome generally show higher expression in female compared to male embryos during early development, similar to X-linked genes (Fig. 2a). There is also a deficit of testis and an excess of ovary expression on the dot, resembling sex-biased expression profiles of the current *Drosophila* X (Fig. 2b). In addition, genes on the dot chromosome are downregulated during male meiosis

(Fig. 2c), mimicking the phenomenon of male germline X inactivation. Finally, flies with only one copy of the dot chromosome are viable and fertile, and the dot is targeted by the chromosome-specific protein Painting of fourth (POF)⁶, which is involved in transcriptional regulation of genes located on the dot chromosome⁶. This resembles the MSL complex, the only other known protein complex that specifically targets a chromosome, and POF may be part of a putative ancestral mechanisms of dosage compensation^{17,18} (Fig. 2d). In some *Drosophila* species, POF shows male-specific binding to the X chromosome in addition to binding to the dot, further supporting its involvement in dosage compensation¹⁷. Some components of the MSL complex are necessary for normal expression of genes located on the dot¹⁹, consistent with interactions of the regulatory network for the current and former sex chromosome of *Drosophila* (Fig. 2e). Thus, many features of the dot chromosome in *Drosophila* resemble unique characteristics of the current X chromosome that distinguish it from autosomes, and can be interpreted as signatures of its former life as a differentiated X chromosome. Female-biased expression during early embryogenesis, an excess of ovary expression and a deficiency of testis expression on Muller F (but not Muller A) are all observed in Diptera species in which the dot segregates as the X (Fig. 3). This confirms that these peculiarities of the dot were present in the X-linked ancestor, and also shows that Muller A only acquired them once it became sex-linked in the lineage leading to *Drosophila*.

It is of interest to consider how this transition could have happened. Sex in most outgroup species of *Drosophila* is determined by the presence of a factor on the Y chromosome (M factor) causing maleness²⁰. Such a dominant-Y system with element F as the sex chromosome could evolve into the *Drosophila* system (dose-dependent sex determination with element A as the sex chromosome) through various intermediate steps¹, and we outline three possible paths that involve mutational events that have been observed in Diptera. This sex-chromosome transition could be initiated by a single epistatic mutation on element A (M*) that makes individuals male regardless of their sex-chromosome karyotype, or through a translocation of the existing male-determining factor (M) onto element A (Fig. 4). Novel sex-determining genes occurring on different chromosomes or translocations of M factors onto autosomes have been observed in houseflies²¹ or humpbacked

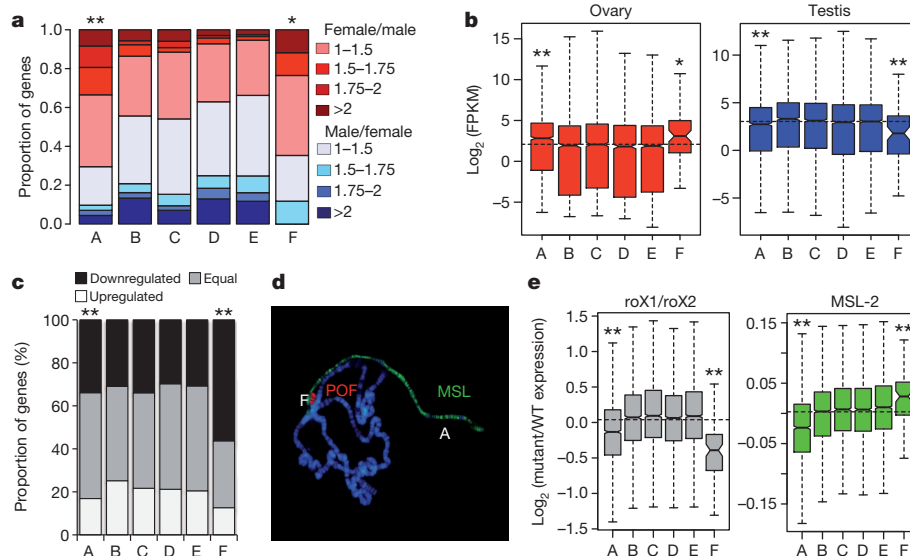


Figure 2 | Properties of the dot chromosome in *Drosophila melanogaster* that resemble that of an X chromosome. **a**, Zygotic transcription, before the onset of MSL-dosage compensation, is female-biased for genes located on the dot (data from ref. 28). **b**, Genes located on the dot chromosome show increased transcription (in fragments per kilobase of transcript per million mapped reads, FPKM) in the ovary and decreased transcription in testis (data from ref. 29). **c**, An excess of genes located on the dot is downregulated during male meiosis

(data from ref. 30). **d**, Chromosome-specific transcriptional regulation of genes located on the dot chromosome by the protein POF. **e**, Misregulation of genes located on the dot chromosome in mutants for some components of the MSL dosage compensation complex (roX1/roX2 mutant in grey; MSL-2 mutant in green; data from ref. 19). Levels of significance (based on resampling for panels **b** and **e** and using Chi-squared tests for panels **a** and **c**) are represented by * $P < 0.05$ and ** $P < 0.01$. Boxes (panels **b** and **e**) as in Fig. 1. WT, wild type.

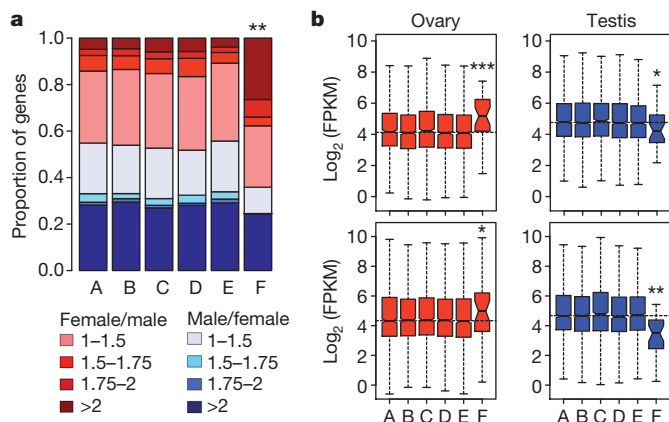


Figure 3 | Gene expression in early embryos and adult gonads in outgroup Diptera species. **a**, Zygotic transcription is female-biased for genes located on the X chromosome (element F) of *B. oleae*, but not the autosomal element A. FPKM values were estimated by RNA-sequencing of male and female stage 5 embryos. **b**, Overexpression in ovary and underexpression in testis on the X chromosome (element F) but not the autosomal element A in *B. oleae* (top) and *S. bullata* (bottom). FPKM values were estimated through RNA-sequencing of dissected testis and ovaries. Levels of significance (based on Chi-square tests for panel **a** and using one-tailed Wilcoxon tests for panel **b**) are represented by * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$. Boxes (panel **b**) as in Fig. 1.

flies²². The fixation of the new male-determining gene on element A would lead to reversal of the ancestral X (Muller F) to an autosome, and the ancestral Y would be completely lost. Loss of the Y is only possible if the ancestral Y chromosome carried no essential male-fertility genes, as seems to be the case in several Diptera species²¹, or if those male-fertility genes moved to another chromosome²³. The

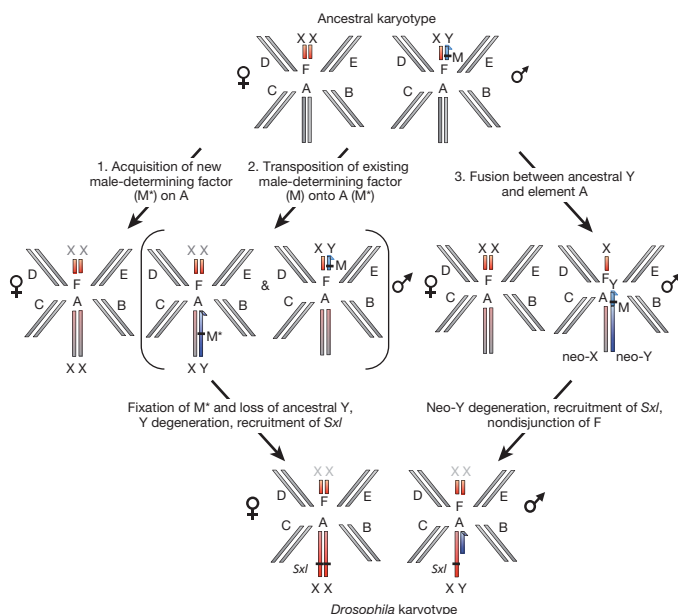


Figure 4 | Turnover of sex chromosomes in *Drosophila*. Hypothetical transition from an ancestral karyotype with element F segregating as a sex chromosome, to the karyotype observed in *Drosophila* in which element A is the sex chromosome. In outgroup species, maleness is determined by the presence of a factor on the Y chromosome (M factor). Element A could either have acquired a new, epistatic M* factor (scenario 1), or the existing M factor could have transposed to element A (scenario 2). This transition could also have been initiated by a fusion of the ancestral Y to element A (scenario 3). Degeneration of the male-limited, non-recombining element A, followed by recruitment of *Sxl* for sex determination, would create the ancestral karyotype of *Drosophila*.

emergence of a male-determining gene on element A would cause male-limited transmission of this chromosome and—as higher Diptera males generally lack recombination²⁴—set in motion genome-wide degeneration of the non-recombining proto-Y. Eventually, *Sxl* was recruited as a dose-dependent sex-determination gene in *Drosophila*, and MSL-mediated dosage compensation evolved on element A¹⁶. The current karyotype of *Drosophila* could also have evolved through a chromosomal fusion between the ancestral Y chromosome and element A (Fig. 4). This would create a male-limited neo-Y (the fused element A) that initially is identical to the neo-X (the unfused element A). The non-recombining neo-Y would undergo chromosome-wide degeneration, and Y-autosome fusions coupled with neo-Y degeneration have happened repeatedly in several *Drosophila* species and other Diptera^{12,25}. This fused neo-Y chromosome would form the current Y of *Drosophila*, and male-fertility genes, if present on the ancestral Y, could be preserved on the chimaeric Y. Both element F and element A would simultaneously segregate as X chromosomes, and eventually *Sxl* would take over the sex-determining function on element A, and a non-disjunction event could restore diploidy for element F in both sexes. Thus, under both the translocation of an existing or the emergence of a new M-factor scenario, the current Y chromosome of *Drosophila* shares no homology to the Y of its ancestor, whereas under the chromosomal-fusion model, the current *Drosophila* Y is a chimaera of the ancestral Y and the degenerated element A and may thus harbour some genes that are also Y-linked in outgroup species. We detect no Y-linked protein-coding genes that are shared between *Drosophila* and Diptera where Muller F is the sex chromosome (Supplementary Tables 3–5 and Supplementary Figs 8–10), providing some support against a fusion between the new and ancestral Y. However, it should be noted that the gene content of the Y is generally poorly conserved, even within *Drosophila*²⁶, and the high turnover of Y-linked genes may obscure any homology between the Y chromosome of drosophilids and that of their outgroups.

Cytological and genetic studies have shown that the dot chromosome of *D. busckii* is fused to both the X and Y chromosome²⁷. Interestingly, this basal species in the *Drosophila* genus also seems to lack MSL-mediated dosage compensation; instead, POF binds to the fused elements A and F in males only¹⁷, and it had been suggested that the A–F fusion represents the ancestral state of *Drosophila*¹⁸. Our data, however, clearly indicate that *D. busckii* harbours a recent, secondary fusion: *P. variegata*, an outgroup species, only has element A as its sex chromosome, and the reduction in male-to-female coverage for element F in *D. busckii* is much less pronounced than for element A. This suggests that reads from the Y-fused homologue of element F are still similar enough to map to the element F portion of the X.

Our whole-genome analysis reveals that heteromorphic sex chromosomes are not necessarily a terminal stage in sex-chromosome evolution. Instead, the dot chromosome of *Drosophila* provides an example of a highly differentiated sex chromosome reverting back to an autosome, and several odd properties of the dot, such as its feminizing effect, its non-random gene content and the presence of chromosome-specific regulation of gene expression, are possible remnants of its former life as a sex chromosome. Such chromosome-specific adaptations of an X chromosome, and male-beneficial genes that may have accumulated on the ancestral Y, probably impose strong fitness constraints that will limit the turnover of highly evolved sex chromosomes in nature.

METHODS SUMMARY

DNA was extracted from one male and one female individual of each species, and 90-base-pair paired-end reads were sequenced with Illumina technology. Genomic reads were trimmed and assembled using SOAPdenovo. Coverage was estimated by aligning reads back to scaffolds using the Burrows-Wheeler Aligner (BWA) and running SoapCoverage on the alignments. Mapping of *D. melanogaster* genes to genomic scaffolds using the BLAST-like alignment tool (BLAT) with a translated query and database allowed assignment of scaffolds to Muller elements. BLAT queries uncovered genes conserved in all species, and their corresponding

protein sequence was extracted using GeneWise and aligned using MUSCLE. After removal of low-quality alignments by Gblocks, the resulting alignments were run through PhyML to obtain the phylogenetic tree and corresponding bootstrap values. Transcriptome data for *D. melanogaster* were obtained from the following sources: male and female early embryo transcriptome data (female-to-male slope expression values for zygotic genes) are taken from supplementary table 1 of ref. 28, ovary and testis expression data were obtained from ref. 29, expression profiles during spermatogenesis are taken from supplementary table 1 of ref. 30, and expression data for MSL mutants and wild type are taken from ref. 19 and were downloaded from the NCBI Gene Expression Omnibus (GEO) depository (GSE3990 and GSE12054). The polytene chromosome image (immunostained for POF and MSL) is a courtesy by J. Larsson. Gene expression in *B. oleae* from early male and female embryos was assayed through RNA-sequencing of stage 5 embryos, following ref. 28, and gene expression profiles for dissected testis and ovaries were obtained from *B. oleae* and *S. bullata*, through RNA-sequencing.

Full Methods and any associated references are available in the online version of the paper.

Received 14 June 2012; accepted 30 April 2013.

Published online 23 June 2013.

- Bull, J. J. *Evolution of Sex Determining Mechanisms* (The Benjamin/Cummings Publishing Company, 1983).
- Sarre, S. D., Ezaz, T. & Georges, A. Transitions between sex-determining systems in reptiles and amphibians. *Annu. Rev. Genomics Hum. Genet.* **12**, 391–406 (2011).
- Pokorná, M. & Kratochvil, L. Phylogeny of sex-determining mechanisms in squamate reptiles: are sex chromosomes an evolutionary trap? *Zool. J. Linn. Soc.* **156**, 168–183 (2009).
- Ohno, S. *Sex Chromosomes and Sex-Linked Genes* (Springer Verlag, 1967).
- Bridges, C. B. Sex in relation to chromosomes and genes. *Am. Nat.* **59**, 127–137 (1925).
- Larsson, J., Chen, J. D., Rasheva, V., Rasmuson-Lestander, A. & Pirrotta, V. Painting of fourth, a chromosome-specific protein in *Drosophila*. *Proc. Natl Acad. Sci. USA* **98**, 6273–6278 (2001).
- Bellott, D. W. & Page, D. C. Reconstructing the evolution of vertebrate sex chromosomes. *Cold Spring Harb. Symp. Quant. Biol.* **74**, 345–353 (2009).
- Muller, H. J. in *The New Systematics* (ed. Huxley, J.) 185–268 (Clarendon Press, 1940).
- Holt, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002).
- Cline, T. W. The *Drosophila* sex determination signal: how do flies count to two? *Trends Genet.* **9**, 385–390 (1993).
- Zacharopoulou, A. Cytogenetic analysis of mitotic and salivary gland chromosomes in the Medfly *Ceratitis capitata*. *Genome* **29**, 67–71 (1987).
- White, M. J. Cytological evidence on the phylogeny and classification of the Diptera. *Evolution* **3**, 252–261 (1949).
- Vicoso, B. & Bachtrög, D. Lack of global dosage compensation in *Schistosoma mansoni*, a female-heterogametic parasite. *Genome Biol. Evol.* **3**, 230–235 (2011).
- Parisi, M. *et al.* Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299**, 697–700 (2003).
- Lifschytz, E. & Lindsley, D. L. The role of X-chromosome inactivation during spermatogenesis. *Proc. Natl Acad. Sci. USA* **69**, 182–186 (1972).
- Gelbart, M. E. & Kuroda, M. I. *Drosophila* dosage compensation: a complex voyage to the X chromosome. *Development* **136**, 1399–1410 (2009).
- Larsson, J., Svensson, M. J., Stenberg, P. & Mäkitalo, M. Painting of fourth in genus *Drosophila* suggests autosome-specific gene regulation. *Proc. Natl Acad. Sci. USA* **101**, 9728–9733 (2004).
- Riddle, N. C., Shaffer, C. D. & Elgin, S. C. A lot about a little dot – lessons learned from *Drosophila melanogaster* chromosome 4. *Biochem. Cell Biol.* **87**, 229–241 (2009).
- Deng, X., Koya, S. K., Kong, Y. & Meller, V. H. Coordinated regulation of heterochromatic genes in *Drosophila melanogaster* males. *Genetics* **182**, 481–491 (2009).
- Sánchez, L. Sex-determining mechanisms in insects. *Int. J. Dev. Biol.* **52**, 837–856 (2008).
- Dübendorfer, A., Hediger, M., Burghardt, G. & Bopp, D. *Musca domestica*, a window on the evolution of sex-determining mechanisms in insects. *Int. J. Dev. Biol.* **46**, 75–79 (2002).
- Traut, W. New Y chromosomes and early stages of sex chromosome differentiation: sex determination in *Megaselia*. *J. Genet.* **89**, 307–313 (2010).
- Larracunte, A. M., Noor, M. A. & Clark, A. G. Translocation of Y-linked genes to the dot chromosome in *Drosophila pseudoobscura*. *Mol. Biol. Evol.* **27**, 1612–1620 (2010).
- Gethmann, R. C. Crossing over in males of Higer Diptera (Brachycera). *J. Hered.* **79**, 344–350 (1988).
- Zhou, Q. & Bachtrög, D. Sex-specific adaptation drives early sex chromosome evolution in *Drosophila*. *Science* **337**, 341–345 (2012).
- Koerich, L. B., Wang, X., Clark, A. G. & Carvalho, A. B. Low conservation of gene content in the *Drosophila* Y chromosome. *Nature* **456**, 949–951 (2008).
- Krivshenko, J. New evidence for the homology of the short euchromatic elements of the X and Y chromosomes of *Drosophila busckii* with the microchromosome of *Drosophila melanogaster*. *Genetics* **44**, 1027–1040 (1959).
- Lott, S. E. *et al.* Noncanonical compensation of zygotic X transcription in early *Drosophila melanogaster* development revealed through single-embryo RNA-seq. *PLoS Biol.* **9**, e1000590 (2011).
- Assis, R., Zhou, Q. & Bachtrög, D. Sex-biased transcriptome evolution in *Drosophila*. *Genome Biol. Evol.* **4**, 1189–1200 (2012).
- Vibrantovski, M. D., Lopes, H. F., Karr, T. L. & Long, M. Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. *PLoS Genet.* **5**, e1000731 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. Kunhi Purayil, I. Kadow, P. Mavragani, J. Larsson and U. Schmidt-Ott for samples, C. Zonneveld, L. Mazzon, D. Obbard, V. Aguiar, N. Gompel and J. Larsson for images, and Q. Zhou and Z. Walton for technical assistance. Funded by National Institutes of Health grants (R01GM076007 and R01GM093182) and a Packard Fellowship to D.B.

Author Contributions B.V. and D.B. designed and performed the study, and wrote the manuscript.

Author Information All the DNA/RNA-sequencing reads generated in this study are deposited in the NCBI Short Reads Archive (<http://www.ncbi.nlm.nih.gov/sra>) under bioprojects SRP021043, SRP021044 and SRP021047. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.B. (dbachtrög@berkeley.edu).

METHODS

Sample collection. *D. busckii* individuals from an inbred isofemale line originally caught in Tallinn, Estonia in the year 2000 were provided by J. Larsson. Wild-caught male and female *P. variegata* were collected in Bari, Italy by S. Kunhi Purayil and I. Kadow, shipped in ethanol and kept at -20°C until DNA extraction. Pupae of *S. bullata* were obtained from Carolina Biological and incubated at 20°C until adults emerged. These were immediately sexed and used for DNA extraction. *B. oleae* males and females were provided by P. Mavragani from a stock maintained at the Aristotle University of Thessaloniki, shipped in ethanol and kept at -20°C until DNA extraction. *H. illucens* were provided by U. Schmidt-Ott, shipped in ethanol and kept at -20°C until DNA extraction.

Library preparation and sequencing. DNA was extracted from a single male and a single female individual from each species using the Puregene Core Kit A (Qiagen). Library preparation and paired-end sequencing of *B. oleae*, *P. variegata*, *S. bullata* and *H. illucens* were performed at the Beijing Genomics Institute (<http://en.genomics.cn/>). The *D. busckii* libraries were prepared in the Bachtrog laboratory following the standard Illumina protocol, and sequenced at the Berkeley sequencing facility. For each species, libraries and sequencing were performed for a single male and a single female separately. The raw sequenced data consists of 90-base pair (bp) paired-end reads, with an insert size of 500 bp for all libraries. All the reads generated for this analysis have been deposited at the NCBI Short Reads Archive under bioproject accession SRP021047.

Genome assembly. We obtained 90-bp paired-end reads for one male and one female of each species, of which the first 10 bp were trimmed. For each species, the male and female genomic reads were assembled using SOAPdenovo (<http://soap.genomics.org.cn/>) with default parameters and a K-mer value of 31. Gaps in the assembly were reduced further using GapCloser (<http://soap.genomics.org.cn/>). Only scaffolds longer than 1,000 bp were kept for further analysis. Statistics for the assemblies are presented in Supplementary Table 1. In the case of *D. busckii*, pooling male and female reads for the assembly led to reduced coverage of element F in both males and females (see below and Supplementary Fig. 2). This is probably due to the hybrid assembly of the neo-X and neo-Y (element F fused to the X, and element F fused to the Y, respectively), as the neo-Y is not yet fully degenerated and some neo-Y-derived reads are included in the assembly. Because this hybrid assembly differs from the true sequence of the neo-X, female reads will not fully map to it, leading to decreased female coverage. This bias disappeared when only female reads were used for the assembly (Supplementary Fig. 3). This female-only assembly was therefore used for all further analyses. It should be noted that this does not affect our conclusions, as the male-to-female coverage ratio is reduced for elements A and F independently of the assembly used in the analysis.

Mapping of the genomic scaffolds to Muller elements. We downloaded all *D. melanogaster* coding sequences from Flybase (<http://www.flybase.org>) and, for each gene, kept only the longest coding sequence. The resulting *D. melanogaster* coding sequences were mapped to the genome assembly of each outgroup using the BLAST-like alignment tool (BLAT)³¹ with both a translated query and a translated database, and only the hit with the highest match score was kept for each gene. Hits with a match score below 50 were excluded. A Perl script was used to count the number of *Drosophila* genes per Muller element that mapped to each scaffold. The scaffold was assigned to the Muller element with the largest number of matching genes. When the same number of genes from different Muller elements mapped to a scaffold, a score was calculated for each Muller element by summing the match scores of all the hits for that element, and the scaffold was assigned to the element with the largest score. The rate of concordance of genes on scaffolds (the number of genes that map to the element their scaffold was assigned to versus total number of genes, calculated using scaffolds carrying at least three genes) is high: *D. busckii*, 96%; *P. variegata*, 92%; *S. bullata*, 80%; *B. oleae*, 88% and *H. illucens*, 71%.

Estimation of male and female genomic coverage. Male and female reads were mapped separately to the genomic scaffolds using the Burrows-Wheeler aligner (BWA)³² set to the default parameters for paired-end reads. The resulting sequence alignment/map (SAM) alignments were used to estimate the male and female coverage depth for each scaffold using SoapCoverage (<http://soap.genomics.org.cn/>). Supplementary Figs 2–7 show the male and female coverage distributions for each Muller element for all the species investigated. For each species, we used one-tailed Wilcoxon tests between each element and the rest of the sample to identify systematic reductions in male-to-female coverage. The results of the tests are presented in Supplementary Table 2.

Building the phylogenetic tree. From the BLAT results (*D. melanogaster* genes versus outgroup genomic sequence, see the section Mapping of the genomic scaffolds to Muller elements above), we selected the 185 *D. melanogaster* genes that had an alignment score above 1,000 for all of the outgroups species. The genomic regions corresponding to this conserved set of genes were extracted from the genomic scaffolds using a Perl script.

GeneWise was used to infer the protein sequence of the genes in the different outgroups from the corresponding genomic region. To avoid potential biases that could arise by using the *D. melanogaster* protein sequences as input for GeneWise (as GeneWise may be more likely to insert errors in the more distant outgroups than in the closer ones), we used *Anopheles gambiae* protein sequences instead, as *A. gambiae* is an outgroup to all the species analysed. For each gene, the inferred protein sequences and the *D. melanogaster* and *A. gambiae* protein sequences were aligned using Muscle. We concatenated the Muscle output files, and ran the resulting concatenated alignment through Gblocks to remove gaps and regions of low alignment quality. The final alignment, consisting of 93,134 amino acids, was used as input for PhyML to obtain the phylogenetic tree and associated bootstrap values (Supplementary Fig. 1).

Resampling procedure. Differences between elements A and F versus autosomes (elements B–E) were tested by resampling n genes 1,000 times from the autosomal sample (where n is the number of genes for element A or F present in the sample). Levels of expression of elements A or F were considered to be significantly different from the autosomes if their observed median fell within the 5% one-sided tail of the resampled distribution.

Expression analysis of adult tissues. *B. oleae* individuals were obtained from olives collected on the University of California, Berkeley campus, and *S. bullata* adults were obtained from pupae purchased from Carolina Biological. Male and female adults were placed in separate vials immediately after emergence, and aged on standard *Drosophila* food for 4 days. Paired-end RNA-sequencing (RNA-seq) reads were obtained from testis and ovary. RNA was extracted from each tissue using a Qiagen RNeasy kit following the manufacturer's protocol. RNA-seq libraries and sequencing were performed at the Beijing Genomics Institute. All the reads generated for this analysis have been deposited at the NCBI Short Reads Archive under bioprojects SRP021043 and SRP021044. For each species, the resulting paired-end 90-bp reads were trimmed, pooled and assembled using SOAPdenovo (<http://soap.genomics.org.cn/soapdenovo.html>) with a K-mer value of 31 and an insert size of 200 bp, and this assembly was further improved using GapCloser (<http://soap.genomics.org.cn/soapdenovo.html>). Only scaffolds longer than 300 bp were kept for further analysis. Scaffolds were mapped to *D. melanogaster* coding sequences downloaded from Flybase (<http://www.flybase.org>) using BLAT; only the location with the best mapping score was kept for each scaffold. When more than one scaffold overlapped on the same *D. melanogaster* gene by more than 20 bp, only the scaffold with the highest mapping score was kept for further analysis. When several scaffolds mapped to different parts of the same gene, or had an overlap shorter than 20 bp, their sequences were concatenated. The resulting *B. oleae* and *S. bullata* gene sequences were assigned to Muller elements according to the genomic location of their *D. melanogaster* homologues. For each adult tissue, reads were mapped to the *B. oleae* or *S. bullata* gene sequences using Bowtie2 (ref. 33) with default parameters. FPKM values were then estimated from the resulting SAM alignments using Cufflinks with default parameters. Genes with FPKM < 10 in both testis and ovary were excluded.

Expression analysis of early embryos. In order to assess early embryonic expression, adult individuals of *B. oleae* were placed in cages with fresh olives. Eggs were periodically collected from the olives and their chorions removed by immersing them in a 50% bleach solution for 1 min. The de-chorionated embryos were staged by observation under a light microscope after immersion in halocarbon oil. Stage 5 embryos were selected, as in *D. melanogaster* they represent the earliest stage containing primarily zygote-derived RNAs (earlier stages contain mostly maternal RNAs²⁸). RNA and DNA were extracted following the protocol described in ref. 28, and the DNA was used to sex the embryos using published sexing primers³⁴. Libraries were made at the Bachtrog laboratory following the protocol of ref. 28 and 100-bp single-end sequencing was performed at the Vincent J. Coates Genomics Sequencing Laboratory, Berkeley. The reads generated for this analysis have been deposited at the NCBI Short Reads Archive under bioproject SRP021044. The first 20 bp of the resulting reads were trimmed before further analysis. Reads were mapped to the *B. oleae* or *S. bullata* gene sequences (obtained in 'Expression analysis of adult tissues') using Bowtie2 (ref. 33) with default parameters. FPKM values were then estimated from the resulting SAM alignments using Cufflinks with default parameters. Genes with FPKM < 1 in both male and female embryos were excluded.

Searching for *Drosophila* Y-linked genes in *Drosophila* outgroups. The male DNA-seq reads available for *P. variegata*, *S. bullata*, *B. oleae* and *H. illucens* were assembled using SOAPdenovo (with a K-mer value of 31), and all scaffolds longer than 200 bp were kept for further analysis. One lane of forward female and male DNA-seq reads were mapped back to these male scaffolds separately using BWA (with default parameters) and male and female coverage was estimated using SoapCoverage. We used single-end mapping of the forward read rather than paired-end mapping because the insert size of the paired-end fragments (500 bp) is longer than the minimum size of the fragments. Similarly, RNA-seq reads obtained

from male tissues were assembled using SOAPdenovo-Trans for *S. bullata* (whole male, male carcass, testis) and *B. oleae* (whole male, male carcass, testis), and all transcripts longer than 200 bp were kept for further analysis. One lane of forward female and male DNA-seq reads were mapped back separately to the resulting male transcripts using BWA (with default parameters) and male and female coverage was estimated using SoapCoverage.

The protein sequences of known ancestral Y-linked genes of *Drosophila*²⁶ (*kl-2*, *kl-3*, *Ory*, *PRY*, *Ppr-Y*, but also *Ccy* and *ARY*, which may be ancestral) were obtained from the NCBI (with identifier (GI) numbers 190608814, 219131049, 16519041, 217416310, 158529626, 190608812 and 281309229, respectively). Genomic scaffolds and transcripts containing homologues to these proteins were identified using tBLASTn (with an *E* value cutoff of 0.01). Sequences identified in our BLAST search that had no coverage in females, but had coverage higher than 0 in males, were classified as candidate homologues to the *Drosophila* Y-linked genes (Supplementary Tables 3 and 4). No such sequences were found in the male transcriptome or genome of *S. bullata*, *B. oleae* and *H. illucens*, the outgroups carrying the ancestral sex chromosome (Muller element F). *P. variegata*, on the other hand, presents strong evidence of sharing both *Ccy* and *kl-2* with other drosophilids. Primers designed to amplify fragments of the candidate Y-linked *P. variegata* scaffolds 13109 and 15584, which share homology with *kl-2* and *Ccy*, respectively, amplified bands of the expected size in male but not female *P. variegata*, confirming their Y linkage (Supplementary Fig. 8).

In Diptera, there is a general tendency for the ribosomal DNA to reside on the sex chromosomes, but its location often differs between closely related species³⁵. For example, both the X and the Y of *D. melanogaster* carry rDNA genes³⁵, and the Y chromosome of *D. pseudoobscura*, which is non-homologous to the Y of *D. melanogaster*, has acquired rDNA loci after its formation²³. On the other hand, rDNA loci in the *Drosophila ananassae* subgroup are located on autosomes, and *D. simulans*, the sister species of *D. melanogaster*, has lost its Y-linked rDNA cluster³⁵. Thus, the absence or presence of rDNA on sex chromosomes contains limited information on whether the Y chromosomes of Diptera are homologous across families, and we therefore focused our analysis on protein-coding genes. Cytological studies have found that some (but not all) true fruitflies have rDNA loci on their Y (including *B. oleae*³⁶), whereas house flies or flesh flies have rDNA clusters on their autosomes^{37,38}.

Candidate Y-linked transcripts of *S. bullata* and *B. oleae* and their *D. melanogaster* homologues. The trimmed whole-body male and female RNA-seq reads of *S. bullata* and *B. oleae*, as well as the male and female DNA-seq reads, were mapped back to the assembled male transcriptome of these species (see section Searching for *Drosophila* Y-linked genes in *Drosophila* outgroups above) using Bowtie2, keeping only reads that mapped fully in the alignment. After filtering out reads with mismatches, the DNA-seq and RNA-seq alignments were used as input for SoapCoverage and Cufflinks, respectively, to estimate the male and female DNA coverage and expression level (in FPKM). As Y-derived sequences are characterized by their lack of female coverage or expression, scaffolds that had at least 6 male reads mapping, had (female reads/male reads) < 0.1 and had (female expression/male expression) < 0.1 were classified as candidate Y-linked transcripts (39 transcripts in *B. oleae* and 87 in *S. bullata*). PCRs with primers designed for 10

candidate transcripts of *S. bullata* yielded male-specific bands for 2 transcripts (standard PCR with an annealing temperature of 58 °C, see Supplementary Fig. 9), confirming that the sample contains Y-derived sequences.

In order to check for Y-linkage of these candidate sequences in *D. melanogaster*, DNA was extracted (using a Qiagen DNeasy kit) from one male and one female *D. melanogaster* and DNA-seq libraries were made and sequenced at Beijing Genomics Institute (paired-end with 500-bp insert size, 2 GB of data per individual). We assembled the resulting trimmed genomic reads with SOAPdenovo (with a K-mer value of 31). The male and female coverage of each genomic scaffold was estimated by mapping the male and female DNA-seq reads against the assembled genome with Bowtie2 (keeping only fully mapped reads, and filtering out reads with mismatches) and processing the SAM alignment with SoapCoverage. The candidate Y-linked transcripts of *S. bullata* and *B. oleae* were then mapped against the *de novo* *D. melanogaster* genome assembly using tBLASTx (with an *E* value cutoff of 0.01, see Supplementary Table 5). As a control, 12 known Y-linked genes of *D. melanogaster* were also mapped to this assembly. All 12 *D. melanogaster* Y-linked genes mapped to scaffolds with male coverage > 0 but no female coverage, confirming that most Y-derived coding sequence is represented in the *de novo* *D. melanogaster* assembly, and can be identified by its patterns of male and female coverage. Contrary to this, only 1 out of 39 of the *B. oleae* and none of the 87 *S. bullata* candidate Y-derived sequences mapped to *D. melanogaster* scaffolds with male-specific coverage, supporting a general lack of homology between the Y-chromosome of drosophilids and that of their outgroups (Supplementary Table 5). Finally, the single *B. oleae* candidate that maps to a male-specific *D. melanogaster* scaffold was shown to yield bands of similar sizes in both sexes when a PCR was performed with primers designed for the scaffold in both *B. oleae* and *D. melanogaster*, showing that this fragment is not Y-linked in either species (Supplementary Fig. 10).

31. Kent, W. J. BLAT – the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
32. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
33. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
34. Gabrieli, P. *et al.* Interchromosomal duplications on the *Bactrocera oleae* Y chromosome imply a distinct evolutionary origin of the sex chromosomes compared to *Drosophila*. *PLoS ONE* **6**, e17747 (2011).
35. Roy, V. *et al.* Evolution of the chromosomal location of rDNA genes in two *Drosophila* species subgroups: *ananassae* and *melanogaster*. *Heredity* **94**, 388–395 (2005).
36. Drosopoulou, E. *et al.* Sex chromosomes and associated rDNA form a heterochromatic network in the polytene nuclei of *Bactrocera oleae* (Diptera: Tephritidae). *Genetica* **140**, 169–180 (2012).
37. Parise-Maltempi, P. P. & Avancini, R. M. P. Cytogenetics of the neotropical flesh fly *Pattonella intermutans* (Diptera, Sarcophagidae). *Genet. Mol. Biol.* **23**, 563–567 (2000).
38. Parise-Maltempi, P. P. & Avancini, R. M. P. Comparative cytogenetic study in Muscidae flies. *Braz. J. Biol.* **67**, 945–950 (2007).

Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex

Jerry L. Chen¹, Stefano Carta^{1,2}, Joana Soldado-Magraner^{1,3}, Bernard L. Schneider⁴ & Fritjof Helmchen^{1,2}

In the mammalian neocortex, segregated processing streams are thought to be important for forming sensory representations of the environment^{1,2}, but how local information in primary sensory cortex is transmitted to other distant cortical areas during behaviour is unclear. Here we show task-dependent activation of distinct, largely non-overlapping long-range projection neurons in the whisker region of primary somatosensory cortex (S1) in awake, behaving mice. Using two-photon calcium imaging, we monitored neuronal activity in anatomically identified S1 neurons projecting to secondary somatosensory (S2) or primary motor (M1) cortex in mice using their whiskers to perform a texture-discrimination task or a task that required them to detect the presence of an object at a certain location. Whisking-related cells were found among S2-projecting (S2P) but not M1-projecting (M1P) neurons. A higher fraction of S2P than M1P neurons showed touch-related responses during texture discrimination, whereas a higher fraction of M1P than S2P neurons showed touch-related responses during the detection task. In both tasks, S2P and M1P neurons could discriminate similarly between trials producing different behavioural decisions. However, in trials producing the same decision, S2P neurons performed better at discriminating texture, whereas M1P neurons were better at discriminating location. Sensory stimulus features alone were not sufficient to elicit these differences, suggesting that selective transmission of S1 information to S2 and M1 is driven by behaviour.

In the whisker region of S1 (also known as the barrel cortex), layer 2/3 (L2/3) pyramidal neurons send direct projections to S2 and M1 in a generally non-overlapping manner^{3–6}. We used a combination of viral-based and synthetic retrograde tracers to label uniquely these long-range projection neurons (Fig. 1a and Supplementary Fig. 1). Using adult transgenic Cre-dependent tdTomato reporter mice⁷, we injected a retrograde-infecting adeno-associated virus (AAV, serotype 6) expressing Cre recombinase (AAV6-Cre)³ into S2 to induce long-term tdTomato expression in S2P neurons, later visualized *in vivo* (Fig. 1b and Supplementary Fig. 2a). M1P neurons were labelled by M1 injection of cholera toxin subunit B (CTB, also known as CtxB) conjugated to Alexa647 (CTB-Alexa647) and retrospectively identified in fixed tissue sections. These tracers exhibit similar relative efficiency, labelling a maximum of approximately 80% of all projection neurons to the injection site (Supplementary Fig. 2b–f). For calcium imaging, we injected into S1 an AAV6 expressing yellow-cameleon Nano140 (YC-Nano140)⁸, a sensitive genetically encoded calcium indicator that reliably reports action-potential firing (Supplementary Fig. 3). Similar to previous reports^{5,6}, we found that S2P and M1P neurons constitute a spatially intermingled, largely non-overlapping population with only a small fraction (approximately 10%) of co-labelled neurons (neurons expressing both S2P and M1P, which we refer to here as ‘S2P/M1P’ neurons). In addition, post-hoc GABA (γ -aminobutyric acid) immunohistochemistry confirmed that labelled S2P and M1P neurons are non-GABAergic (Supplementary Fig. 4).

Mice use their whiskers to discriminate fine tactile features and to determine an object’s position¹. We first measured neuronal firing of long-range projection neurons in mice trained to perform an S1-dependent texture discrimination task^{9–12} under head-fixed conditions, enabling simultaneous two-photon imaging (Fig. 1c, d). Using a ‘go/no-go’ behaviour paradigm, mice were trained for ‘go’ trials to lick for a water reward (‘hit’) when presented with a target texture, a panel of coarse sandpaper (P100); and for ‘no-go’ trials to withhold licking (‘correct rejection’) when presented with one of three non-rewarded, ‘non-target’ textures of increasingly smoother grades (P280, P600, P1200) (Supplementary Fig. 5 and Supplementary Videos 1 and 2). Misses on go trials were not rewarded, and false alarms on no-go trials were punished with an air puff and a time-out period.

Calcium imaging of S1 projection neurons was carried out in six expert mice while whisker motion and texture contact were monitored using high-speed videography (Fig. 1e and Supplementary Video 3). We measured neuronal activity across tens of labelled cells, sampling different neurons across imaging sessions. We analysed 679 neurons, including 161 S2P, 118 M1P, 31 S2P/M1P, and 369 unlabelled neurons (for S2P/M1P neuron analysis, see Supplementary Table 1). Unlabelled neurons probably consist of a mixed population of unlabelled S2P or M1P neurons, and neurons not projecting to these areas. Responses in S1 were sparse, with 34% of L2/3 neurons showing activity during the behaviour session. S1 neurons are responsive to both whisker motion during free whisking^{13,14} and sensory input from whisker touch^{15–17}. Our image acquisition rate was too low to observe locking of neuronal responses with rhythmic whisking¹⁵ or surface palpitations⁹, so for a simple functional classification we cross-correlated cellular calcium signals across all trials against the envelope of whisking amplitude and a binary vector representing touch periods, respectively (Fig. 1f, g). Based on positive correlation peaks within defined lag time windows and significant differences in the two correlations, cells were classified as ‘whisking’ or ‘touch’ cells; cells remained ‘unclassified’ when no significant positive correlation to whisking or touch existed or correlation peaks occurred outside of the lag windows (see Methods and Supplementary Fig. 6). A large fraction of active S2P neurons (47%) were responsive to touch, higher compared to M1P and unlabelled neurons ($P < 0.05$, bootstrap test; Fig. 1h). Notably, whisking cells were identified among S2P and unlabelled neurons but not among M1P neurons, suggesting that whisking information arriving in S1 from the sensory periphery¹⁴ or from M1 (refs 18, 19) is transmitted to S2 but not to M1 (ref. 20). Estimates of neuronal firing from calcium signals (Supplementary Fig. 3e) indicate that unclassified neurons were less active than touch or whisking neurons (Supplementary Fig. 7a–c). It is unclear whether unclassified neurons are relevant to task behaviour or simply exhibit spontaneous activity.

We next examined how S2P or M1P neurons might contribute to sensory coding or decision making during texture discrimination. Individual neurons exhibited diverse responses and selectivity to trial type and texture (Fig. 2a, b). To quantify how well single cells could

¹Brain Research Institute, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. ²Neuroscience Center Zurich, University of Zurich/ETH Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. ³Institute of Neuroinformatics, University of Zurich/ETH Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. ⁴Brain Mind Institute, Ecole Polytechnique Fédérale de Lausanne (EPFL), EPFL SV BMI LEN, Station 19, CH-1015 Lausanne, Switzerland.

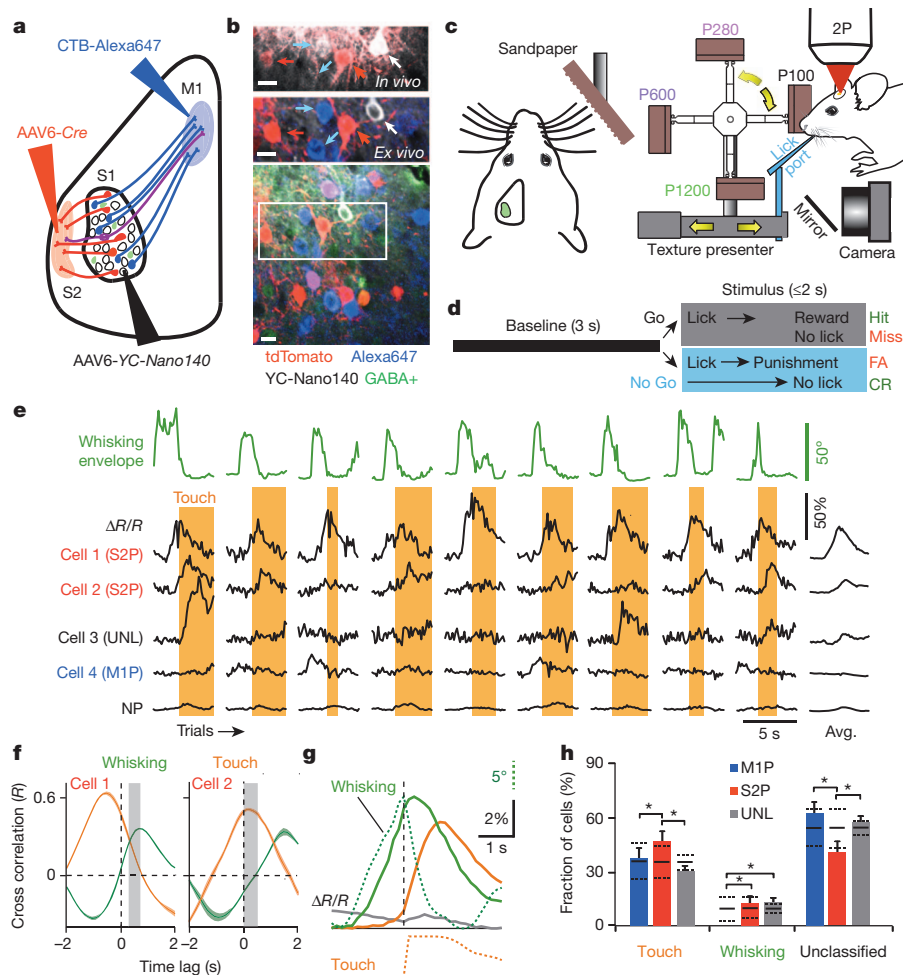


Figure 1 | *In vivo* calcium imaging of long-range projection neurons in S1 during texture discrimination. **a**, Retrograde labelling of S2P and M1P neurons in S1 by AAV6-Cre and CTB-Alexa647 injection, respectively. AAV6-YC-Nano140 was injected into S1 for calcium imaging. **b**, Top, *in vivo* image of L2/3 neurons in S1 expressing YC-Nano140 (white). AAV6-Cre-infected S2P neurons express tdTomato (red). Middle and bottom (lower magnification), post-hoc identification of *in vivo* imaged neurons with CTB-Alexa647-labelled M1P neurons (blue) and immunostained GABA-positive neurons (green). Scale bar, 20 μ m. **c**, Setup for two-photon (2P) imaging of S1 neurons during head-fixed texture discrimination. **d**, Trial structure for go/no-go texture discrimination task. CR, correct rejection; FA, false alarm. **e**, Calcium transients (black) from example cells and the neuropil (NP) across nine trials,

discriminate between any given two conditions, between the reported decision of hit versus correct rejection or between two distinct textures, we performed a receiver operating characteristic (ROC) analysis^{16,21} (Fig. 2c; see Supplementary Fig. 8 for neuronal populations). When comparing hit and correct rejection trials, we found that 49% of S2P neurons discriminated these trial types above chance, a higher percentage than for M1P or unlabelled neurons ($P < 0.05$, bootstrap test; Fig. 2d). For analysis of texture coding, we focused only on pairs of non-target textures (non-rewarded stimuli, to which animals should respond by withholding licking) in correct rejection trials to control for behavioural responses. For each non-target texture pair, a larger fraction of active M1P neurons showed above-chance discriminative power compared to S2P neurons ($P < 0.05$, bootstrap test; Fig. 2d). However, although fewer in numbers, individual discriminative S2P neurons were more accurate than M1P neurons in discriminating P280 versus P600, and P280 versus P1200 textures ($P < 0.05$, one-way analysis of variance (ANOVA), Tukey's post-hoc test; Fig. 2e). Although discriminating cells were found in each behavioural category, touch-related neurons performed better than whisking or unclassified

and the average trace across all trials. Envelope of whisking amplitude (green) and periods of touch (orange area) are also shown. UNL, unlabelled. **f**, Cross-correlation analysis of calcium signals with whisking amplitude (green) and touch (orange) across different time lags for 'whisking' (cell 1) and 'touch' (cell 2) in **e**. Shaded trace indicates 95% confidence interval from bootstrap test. Grey area indicates lag window for classification. **g**, Average calcium trace across all whisking (green), touch (orange), and unclassified (grey) neurons shown with average whisking amplitude and touch vectors (dotted lines). **h**, Distribution of classified cells across subtypes. Error bars, s.d. from bootstrap test. A permutation test of shuffled labels are shown. Solid horizontal lines, means; dashed horizontal lines, 95% confidence intervals ($n = 231$ active neurons). $*P < 0.05$.

cells. We conclude that differences exist between M1P and S2P neurons with respect to the information available for texture coding and decision making.

We next asked whether these differences are intrinsic to these cell types or whether they are influenced by the sensory-processing requirements specific to this texture-discrimination task. We reasoned that a contrasting task with different requirements could result in a different set of activation patterns between M1P and S2P neurons. To address this, we trained a second set of mice ($n = 4$) to a variation of a head-fixed, go/no-go object-localization task (also known as 'object detection with distractors'²²) previously reported to activate cortical circuits between S1 and M1 (refs 16, 18, 19, 23, 24) (Fig. 3a, b). In this task, mice are trained to use multiple whiskers to locate the same vertical pole presented at different positions along the anterior-posterior axis. The pole was presented at a fixed target location for go trials and at two non-target locations (+4.29 and +6 mm) anterior to the target location for no-go trials. We used an identical trial structure to the texture-discrimination task to control for non-sensory differences across tasks. Mice could reach similar performance levels²² compared to texture

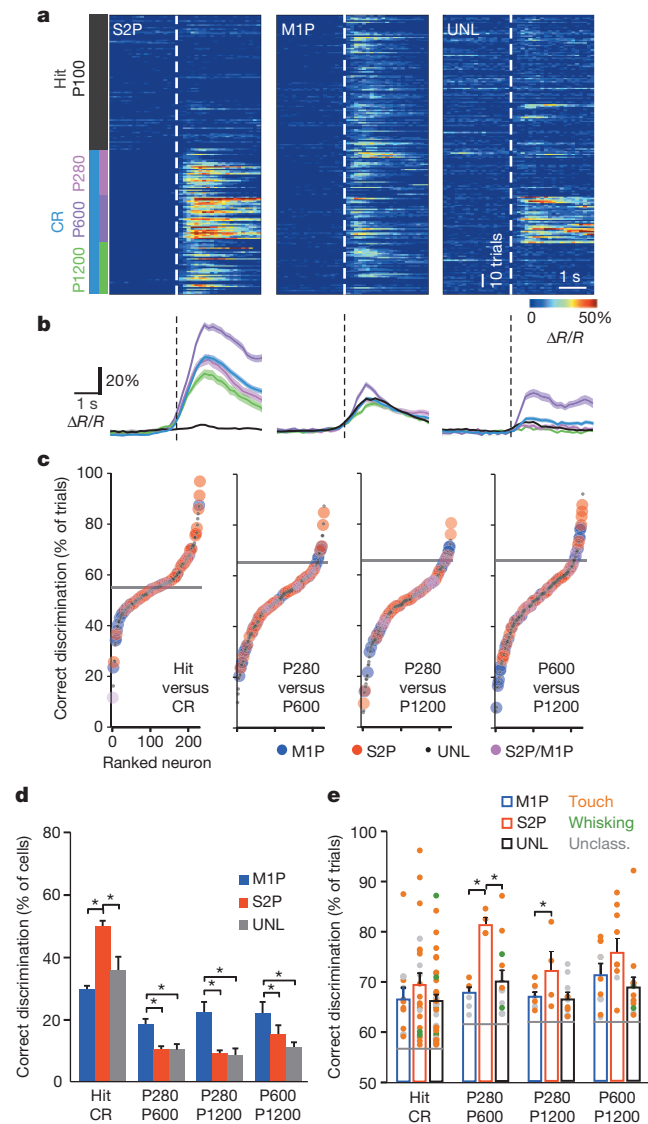


Figure 2 | Single-neuron discrimination analysis of decision or texture in S1 projection neurons. **a**, Single-trial responses of example S2P, M1P or UNL touch neurons to trial type or texture aligned to first touch (dotted line). **b**, Average calcium transient of neurons in **a** according to trial type or texture. Shaded areas, s.e.m. **c**, Fraction of trials in which individual cells correctly discriminated between decision (hit versus correct rejection) or between non-target textures from ROC analysis. Grey line indicates the 95th percentile of distribution from a permutation test of decision or texture labels. Neurons are ranked according to the fraction of trials that were correctly discriminated. Neurons above this line can discriminate above chance. **d**, Fraction of active cells discriminating trial type or texture above chance. **e**, Performance of neurons discriminating above chance. Circles indicate individual neurons shaded according to their behaviour classification. Grey lines indicate 95th percentile of distribution from a permutation test of decision or texture labels. Error bars, s.d. from a permutation test (**d**), s.e.m. (**e**). $n = 231$ active neurons; $*P < 0.05$.

discrimination (Supplementary Fig. 9). Calcium imaging in 386 neurons (84 M1P, 85 S2P, 19 S2P/M1P, and 198 unlabelled neurons) revealed a higher fraction of active cells (47%) compared to during texture discrimination. A large fraction of active M1P neurons (54%) were responsive to touch with a reduced representation (18%) in S2P neurons ($P < 0.05$, bootstrap test; Fig. 3c, d and Supplementary Fig. 7d–f). Similar to texture discrimination, no whisking-related M1P neurons were identified during object localization, demonstrating that this differential motor representation across projection neurons is intrinsic to the circuit. A ROC analysis revealed that similar fractions of S2P and

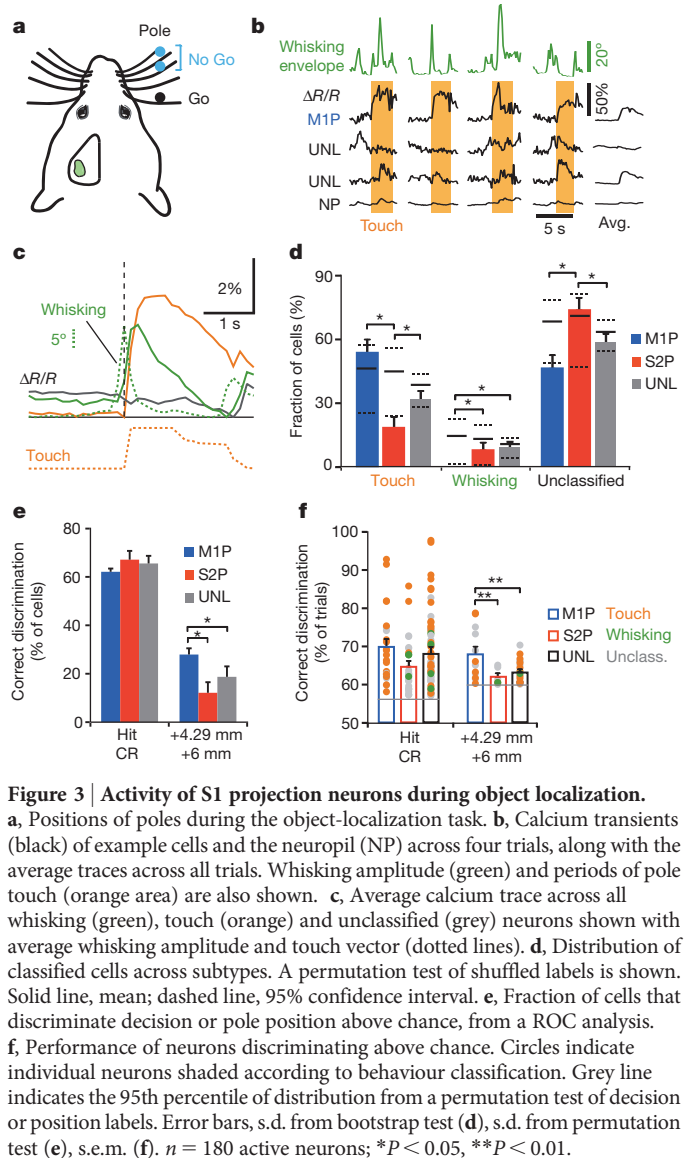


Figure 3 | Activity of S1 projection neurons during object localization.

a, Positions of poles during the object-localization task. **b**, Calcium transients (black) of example cells and the neuropil (NP) across four trials, along with the average traces across all trials. Whisking amplitude (green) and periods of pole touch (orange area) are also shown. **c**, Average calcium trace across all whisking (green), touch (orange) and unclassified (grey) neurons shown with average whisking amplitude and touch vector (dotted lines). **d**, Distribution of classified cells across subtypes. A permutation test of shuffled labels is shown. Solid line, mean; dashed line, 95% confidence interval. **e**, Fraction of cells that discriminate decision or pole position above chance, from a ROC analysis. **f**, Performance of neurons discriminating above chance. Circles indicate individual neurons shaded according to behaviour classification. Grey line indicates the 95th percentile of distribution from a permutation test of decision or position labels. Error bars, s.d. from bootstrap test (**d**), s.d. from permutation test (**e**), s.e.m. (**f**). $n = 180$ active neurons; $*P < 0.05$, $**P < 0.01$.

M1P neurons could discriminate hit and correct rejection trials with similar performance levels (Fig. 3e, f and Supplementary Fig. 10). However, a larger fraction of M1P neurons could discriminate pole position at +4.29 mm from +6 mm compared to both S2P and unlabelled neurons ($P < 0.05$, bootstrap test) and with better accuracy ($P < 0.01$, one-way ANOVA, Tukey's post-hoc test; for touch neurons only see Supplementary Fig. 11).

The results suggest that, depending on task conditions, M1P and S2P neurons may be recruited in a manner necessary to perform sensory-driven, goal-directed behaviour. However, cell-type differences could arise simply from differences in physical stimuli used in each task. Indeed, fine-scale analysis of principal whisker kinematics showed that sandpaper and pole contacts showed distinct kinematic features. Changes in whisker curvature that relate to contact forces²² as well as the frequency of high acceleration-velocity 'slip' events associated with texture coding^{11,12} were greater during texture discrimination compared to object localization ($P < 0.001$, Kolmogorov–Smirnov test; Fig. 4a and Supplementary Fig. 12a). During object localization, pole contact by the principal whisker occurred along a greater range of angles ($P < 0.001$, F-test). In addition, contact occurred occasionally with different sets of whiskers across pole positions (Supplementary Fig. 13a, b). This pattern of whisker contact suggests that haptic sensing and labelled-line encoding strategies²⁴ are both available to the animal

in this task. We next asked whether S2P and M1P neurons show differential sensitivity to these parameters across tasks. Cross correlation with calcium signals of touch cells on correct rejection trials revealed diverse and sometimes high correlation with kinematic features (Fig. 4b, see also Supplementary Figs 12b and 13c, d). However, correlation of activity in S2P and M1P neurons to curvature change or

mean touch angle were not different during both tasks. Although M1P neurons displayed higher correlation to slip events than S2P neurons, this difference was present during both task conditions (sandpaper, $P < 0.05$; pole, $P < 0.02$, one-way ANOVA, Tukey's post-hoc test; Fig. 4b). Thus, although differences exist in whisker kinematics, the cellular responses related to these parameters do not differ across cell types in a task-dependent manner.

Given our limitations in temporal resolution to account for all aspects of touch-related neuronal activity during the complex, multi-whisker interactions, we also measured the responses of S2P and M1P neurons in naive, non-rewarded animals presented passively with both sandpaper and pole under simulated task conditions ($n = 5$ mice; 76 M1P, 58 S2P, 2 S2P/M1P, and 352 unlabelled neurons; summarized in Supplementary Fig. 7f). Similar fractions of active S2P and M1P neurons were responsive to touch for sandpaper and pole presentation (Fig. 4c). We found a diverse set of responses with certain neurons responding to either one or both stimuli but no difference in bias between S2P and M1P neurons (Fig. 4d). A ROC analysis showed a substantial reduction in the fraction of cells discriminating target versus non-target stimuli in naive animals, when compared to hit versus correct rejection trials in trained animals during texture discrimination and, to a lesser degree, during object localization ($P < 0.05$, bootstrap test; Fig. 4e and Supplementary Fig. 14). When comparing non-target stimuli, the fractions of discriminative cells were similar across cell types, which is not the case in task-performing animals. These results demonstrate further that sensory stimulus features alone are not sufficient to explain the different activation patterns observed in S2P and M1P neurons under distinct task conditions.

In conclusion, we show that activity in S1 can be routed selectively to different cortical areas based on the sensory processing requirements for the execution of specific tasks. The prominent recruitment of S2P neurons during texture discrimination compared to that during object localization may reflect the activation of higher sensory areas² for processing more complex sensory stimuli^{25,26}. In contrast, M1P neurons show an increased response to, and ability to discriminate, pole position during object localization. Although the behaviour strategy used by the animal ('active sensing' versus 'detection amid distractors')^{22,24} cannot be distinguished, our activation patterns are in line with previous reports indicating that S1–M1 integration may inform the decision of the mouse under this task condition^{18–20}. As sensory stimulus features alone are not sufficient to produce these differential activation patterns among S2P and M1P neurons, we speculate that other mechanisms could be involved, including plasticity of local and long-range circuits during task learning^{27,28} or top-down influences exerted by feedback circuits or attention-related brain areas during task engagement^{29,30}. Understanding the circuits and mechanisms underlying this selective routing of sensory information will warrant further investigation.

METHODS SUMMARY

Young adult male transgenic Cre-dependent tdTomato reporter mice (P35–42) were injected with virus expressing YC-Nano140 into S1 and retrograde virus expressing Cre into S2. A cranial window was implanted over S1, along with a head post for head fixation. After a recovery period of 1 week, animals were habituated to head fixation and trained in the texture-discrimination or object-localization task, respectively. In separate experiments, sandpaper or a pole were presented to naive animals under similar task conditions but with reward and punishment removed. The principal whisker corresponding to imaging areas in YC-Nano140 expression regions was identified by intrinsic signal optical imaging. *In vivo* two-photon calcium imaging (7 Hz) in S1, along with high-speed videography (500 Hz) of motion and touch of all contralateral whiskers was carried out on animals, 3 to 4 weeks post injection. Trained mice were imaged twice a day for up to 16 behaviour sessions. Untrained mice were imaged for one session each of sandpaper and pole presentation. Animals were injected with CTB-Alexa647 into M1 after the last imaging session, then perfused and fixed after 5 to 7 days with 4% paraformaldehyde. GABA immunohistochemistry was carried out on fixed tissue sections cut parallel to the imaging plane. *In vivo* imaged areas were located in post-hoc sections to identify M1P neurons and GABAergic interneurons. Data analysis

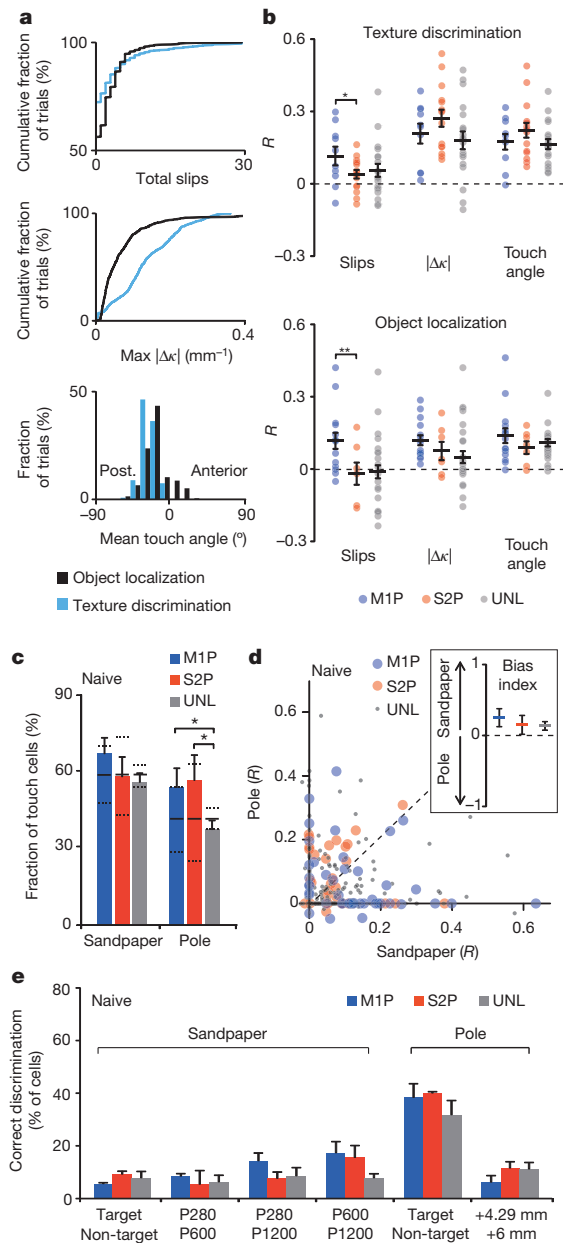


Figure 4 | Sensory stimuli are not sufficient to produce task-related differences. **a**, Whisker kinematic differences during texture discrimination versus object localization for high-acceleration-velocity 'slip' events, maximum absolute curvature change ($\max |\Delta\kappa|$), mean touch angle (0° = orthogonal to anterior–posterior axis) over first second of touch. **b**, Cross correlation of calcium signals with whisker kinematic features for touch during texture discrimination (top panel) or object localization (bottom panel). **c**, Fraction of touch neurons identified by cross-correlation analysis from active neurons in naive animals. A permutation test of shuffled labels is shown. **d**, Cross correlation of calcium activity to sandpaper versus pole touch in naive animals. Inset, population bias index (see Methods) determined from the R value. **e**, Fraction of cells discriminating above chance for target versus non-target stimuli or between non-target stimuli, from a ROC analysis in naive animals. Error bars, s.e.m. (**a**), s.d. from a bootstrap test (**c**), s.d. from a permutation test (**c**). $n = 1,574$ trials (**a**), $n = 98$ touch neurons (**b**); $n = 207$ active neurons (**c–e**); * $P < 0.05$, ** $P < 0.02$.

was performed with ImageJ and MATLAB. Behaviour and whisker kinematic vectors were obtained by analysis of high-speed videos of whisker movement. Cells were classified according to behaviour using cross-correlation analysis of calcium signals and behaviour vectors. Single-cell discrimination of decision or stimulus was performed using a ROC analysis.

Full Methods and any associated references are available in the online version of the paper.

Received 27 September 2012; accepted 30 April 2013.

Published online 23 June 2013.

- Diamond, M. E., von Heimendahl, M., Knutsen, P. M., Kleinfeld, D. & Ahissar, E. 'Where' and 'what' in the whisker sensorimotor system. *Nature Rev. Neurosci.* **9**, 601–612 (2008).
- Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
- Aronoff, R. et al. Long-range connectivity of mouse primary somatosensory barrel cortex. *Eur. J. Neurosci.* **31**, 2221–2233 (2010).
- Mao, T. et al. Long-range neuronal circuits underlying the interaction between sensory and motor cortex. *Neuron* **72**, 111–123 (2011).
- Sato, T. R. & Svoboda, K. The functional properties of barrel cortex neurons projecting to the primary motor cortex. *J. Neurosci.* **30**, 4256–4260 (2010).
- Chakrabarti, S. & Alloway, K. D. Differential origin of projections from SI barrel cortex to the whisker representations in SII and M1. *J. Comp. Neurol.* **498**, 624–636 (2006).
- Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
- Horikawa, K. et al. Spontaneous network activity visualized by ultrasensitive Ca^{2+} indicators, yellow Cameleon-Nano. *Nature Methods* **7**, 729–732 (2010).
- Carvell, G. E. & Simons, D. J. Biometric analyses of vibrissal tactile discrimination in the rat. *J. Neurosci.* **10**, 2638–2648 (1990).
- Guic-Robles, E., Jenkins, W. M. & Bravo, H. Vibrissal roughness discrimination is barrel cortex-dependent. *Behav. Brain Res.* **48**, 145–152 (1992).
- Jadhav, S. P., Wolfe, J. & Feldman, D. E. Sparse temporal coding of elementary tactile features during active whisker sensation. *Nature Neurosci.* **12**, 792–800 (2009).
- von Heimendahl, M., Itskov, P. M., Arabzadeh, E. & Diamond, M. E. Neuronal activity in rat barrel cortex underlying texture discrimination. *PLoS Biol.* **5**, e305 (2007).
- Crochet, S. & Petersen, C. C. Correlating whisker behavior with membrane potential in barrel cortex of awake mice. *Nature Neurosci.* **9**, 608–610 (2006).
- Fee, M. S., Mitra, P. P. & Kleinfeld, D. Central versus peripheral determinants of patterned spike activity in rat vibrissa cortex during whisking. *J. Neurophysiol.* **78**, 1144–1149 (1997).
- Curtis, J. C. & Kleinfeld, D. Phase-to-rate transformations encode touch in cortical neurons of a scanning sensorimotor system. *Nature Neurosci.* **12**, 492–501 (2009).
- O'Connor, D. H., Peron, S. P., Huber, D. & Svoboda, K. Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron* **67**, 1048–1061 (2010).
- de Kock, C. P., Bruno, R. M., Spors, H. & Sakmann, B. Layer- and cell-type-specific suprathreshold stimulus representation in rat primary somatosensory cortex. *J. Physiol. (Lond.)* **581**, 139–154 (2007).
- Petreaanu, L. et al. Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* **489**, 299–303 (2012).
- Xu, N. L. et al. Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature* **492**, 247–251 (2012).
- Hill, D. N., Curtis, J. C., Moore, J. D. & Kleinfeld, D. Primary motor cortex reports efferent control of vibrissa motion on multiple timescales. *Neuron* **72**, 344–356 (2011).
- Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, 1966).
- O'Connor, D. H. et al. Vibrissa-based object localization in head-fixed mice. *J. Neurosci.* **30**, 1947–1967 (2010).
- Huber, D. et al. Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature* **484**, 473–478 (2012).
- Mehta, S. B., Whitmer, D., Figueroa, R., Williams, B. A. & Kleinfeld, D. Active spatial perception in the vibrissa scanning sensorimotor system. *PLoS Biol.* **5**, e15 (2007).
- Melzer, P., Champney, G. C., Maguire, M. J. & Ebner, F. F. Rate code and temporal code for frequency of whisker stimulation in rat primary and secondary somatic sensory cortex. *Exp. Brain Res.* **172**, 370–386 (2006).
- Alloway, K. D. Information processing streams in rodent barrel cortex: the differential functions of barrel and septal circuits. *Cereb. Cortex* **18**, 979–989 (2008).
- Guic, E., Carrasco, X., Rodriguez, E., Robles, I. & Merzenich, M. M. Plasticity in primary somatosensory cortex resulting from environmentally enriched stimulation and sensory discrimination training. *Biol. Res.* **41**, 425–437 (2008).
- Wiest, M. C., Thomson, E., Pantoja, J. & Nicolelis, M. A. Changes in S1 neural responses during tactile discrimination learning. *J. Neurophysiol.* **104**, 300–312 (2010).
- Krupa, D. J., Wiest, M. C., Shuler, M. G., Laubach, M. & Nicolelis, M. A. Layer-specific somatosensory cortical activation during active tactile discrimination. *Science* **304**, 1989–1992 (2004).
- Gilbert, C. D. & Sigman, M. Brain states: top-down influences in sensory processing. *Neuron* **54**, 677–696 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. Soldado-Magraner and L. Sumanovski for assistance with data analysis, H. Kasper, M. Wieckhorst, S. Giger and F. Voigt for technical assistance, A. Miyawaki for plasmid reagents, and D. Margolis, H. Lütcke and K. Schulz for help with initial experiments, helpful discussions and comments on the manuscript and V. Padrun and F. Pidoux for virus production. This work was supported by grants from the Swiss National Science Foundation (310030-127091 to F.H.), the EU-FP7 program (PLASTICISE project 223524 to F.H. and B.L.S.; and the BRAIN-I-NETS project 243914 to F.H.), the Swiss SystemsX.ch initiative (project 2008/2011-Neurochoice to F.H. and B.L.S.), the National Center of Competence in Research 'Neural Plasticity and Repair' (F.H.), Forschungskredit of the University of Zurich (grant 541541808 to J.L.C.) and a fellowship from the US National Science Foundation, International Research Fellowship Program (grant 1158914 to J.L.C.).

Author Contributions J.L.C. and F.H. designed the study. J.L.C. carried out experiments. J.L.C., S.C., J.S.M. and F.H. performed data analysis. S.C. carried out experiments and data analysis characterizing YC-Nano140. B.L.S. contributed viral reagents. J.L.C. and F.H. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to F.H. (helmchen@hifo.uzh.ch).

METHODS

Generation of viral construct. For the construction of the *pAAV-EF1 α -YC-Nano140* viral construct, a *YC-Nano140* insert with 5' BamHI and 3' EcoRI restriction sites was generated by polymerase chain reaction (PCR) amplification from a *pcDNA3-YC-Nano140* vector⁸ and subcloned into an *pAAV-EF1 α -dIO-eYFP* plasmid³¹. The *pAAV-pgk-Cre* construct has been described previously³. Recombinant serotype 6 AAV particles were produced by co-transfecting AAV-293 cells with the shuttle plasmid and the *pDP6* packaging plasmid. Cell lysates were subjected to purification on iodixanol density gradients followed by HPLC with HiTrap Heparin column (GE Healthcare Bio-Sciences AB) using standard procedures. The viral suspension that was obtained was concentrated using Centricon centrifugal filter devices with a molecular weight cut-off of 100 kilodaltons (kDa) (Millipore), and the suspension medium was replaced with PBS. Vector titres were determined by measuring the number of encapsidated genomes per ml using real-time PCR.

Viral and tracer injections. Experimental procedures were carried out following the guidelines of the Veterinary Office of Switzerland and were approved by the Cantonal Veterinary Office in Zurich. Stereotaxic viral and tracer injections were performed on young adult (postnatal day 35 to 42) male transgenic Cre-dependent tdTomato reporter mice (*ROSA26tm14(CAG-tdTomato)*, Ai14) as described previously³². *AAV6-EF1 α -YC-Nano140* (300 nl, approximately 1×10^9 vg μ l⁻¹) was delivered into S1, targeting L2/3 (1.1 mm posterior to bregma, 3.3 mm lateral, approximately 300 μ m below the pial surface). *AAV6-pgk-Cre* (300 nl, approximately 1×10^9 vg μ l⁻¹) was delivered into S2, targeting L2/3 and L5 (0.7 mm posterior to bregma, 4.2 mm lateral, approximately 300 and 500 μ m below the pial surface). CTB-Alexa647 (Molecular Probes, Invitrogen; 300 nl, 1% wt/vol) was injected into M1 immediately after the last behaviour imaging session, targeting L2/3 and L5 (1.2 mm anterior to bregma, 0.6 mm lateral, approximately 300 and 500 μ m below the pial surface).

Cranial window implantation and habituation. To allow long-term *in vivo* calcium imaging, a cranial window was implanted 24 h after virus injections over S1 as described³³. A metal post for head fixation was implanted on the skull, contralateral to the cranial window, using dental acrylic. One week after chronic window implantation, mice were handled daily for 1 week while they became acclimatized to a minimum of 15 min of head fixation. Mice were water restricted for the remainder of the experiment.

Texture-discrimination and object-localization tasks. Behaviour experiments were performed using a data acquisition interface (USB-6008; National Instruments) and custom-written LabVIEW software (National Instruments) to control devices required for the task and for recording trial and licking data. Licking of a water port mounted to a piezo film sensor (MSP1006-ND; Measurement Specialties) triggered delivery of water (5 to 6 μ l) through a miniature rocker solenoid valve (0127; Buerkert). Mice were initially trained during two sessions to trigger water reliably by licking the lick port, and then progressed to go/no-go task training. For texture discrimination, commercial-grade sandpaper (3M) was used. Rough sandpaper (P100) served as a target stimulus and sandpapers of increasing smoothness (P280, P600, P1200) served as non-target stimuli. Sandpapers were mounted onto panels attached to a stepper motor (T-NM17A04; Zaber) mounted onto a motorized linear stage (T-LSM100A; Zaber) to move textures in and out of reach of whiskers. For object localization, a 1.6-mm diameter pole was mounted onto a pneumatic linear slider (SLS-10-30-P-A Mini slide; Festo) to bring the pole in and out of reach of the whiskers rapidly²². This device was mounted on the same motorized linear stage used for texture discrimination to position the pole at target and non-target locations (4.29 mm or 6 mm anterior to the target location), all within reach of the whiskers. Go and no-go trials were presented randomly with a 50% probability of each trial type, and with a maximum of three consecutive presentations of the same trial type. A trial consisted of a 3-s pre-stimulus period followed by stimulus presentation for a maximum of 2 s accompanied by an intermittent 2,093-Hz auditory cue tone. Licking during target stimulus presentation was scored as a 'hit' and triggered immediate withdrawal of the stimulus accompanied by water delivery paired with a 2,093- and 2,793-Hz auditory reward tone. When there was no lick, this was scored as a 'miss', and the next trial followed immediately. During presentation of non-target stimuli, a no-lick response was scored as a 'correct rejection', and the next trial followed immediately; and licking was scored as a 'false alarm', in which case the stimulus was withdrawn immediately, no water reward was given, and the animal was punished with a light air puff to the face and a 7-s time-out period accompanied by auditory white noise. A sensitivity index of animal performance (d') was calculated for all trials per session as $d' = Z(\text{Hit}/(\text{Hit} + \text{Miss})) - Z(\text{FA}/(\text{FA} + \text{CR}))$, where Hit and Miss are the numbers of hits and misses, respectively, CR is the number of correct rejections, FA is the number of false alarms. $Z(p)$, $p \in [0, 1]$ is the inverse of the cumulative Gaussian distribution, where p is Hit/(Hit + Miss) or FA/(FA + CR). Imaging during behaviour began once animals reached a performance level of $d' > 1.75$

(80% correct) for one session. All training and imaging sessions occurred twice per day. For passive presentation of sandpaper or pole in untrained animals, a similar trial structure was used as during task conditions, with the lick port removed and no reward or punishment given. For both trained and untrained animals, stimuli were presented at positions in which whisker contact could occur under both non-whisking and whisking conditions.

Intrinsic signal optical imaging. The S1 barrel column corresponding to the *YC-Nano140* expression area was identified using intrinsic signal optical imaging under approximately 1.5% isoflurane anaesthesia. The cortical surface was illuminated with a 630-nm light-emitting diode (LED), single whiskers were stimulated (2 to 4° rostrocaudal deflections at 10 Hz), and reflectance images were collected through a $\times 4$ objective with a CCD camera (Toshiba Teli CS3960DCL; 12-bit; 3-pixel binning, 427×347 binned pixels, 8.6- μ m pixel size, 10-Hz frame rate). Intrinsic signal changes were computed as fractional changes in reflectance relative to the pre-stimulus average (50 frames; expressed as $\Delta R/R$). The centres of the barrel columns corresponding to stimulated whiskers were located by averaging intrinsic signals (15 trials), median-filtering (5-pixel radius) and thresholding to find signal minima. Reference surface vasculature images were obtained using 546-nm LED and matched to images acquired during two-photon imaging. All whiskers, with the exception of principal and first-order surround whiskers corresponding to the expression area, were partially trimmed prior to behaviour training, to a length at which they were out of reach from pole or texture contact during the task. During whisker trimming, the principal whisker was noted by images taken from the high-speed video camera for re-identification in subsequent imaging sessions for whisker tracking.

Two-photon imaging. We used a custom-built two-photon microscope controlled by HelioScan³⁴, equipped with a Ti:sapphire laser system (approximately 100-femtosecond (fs) laser pulses; Mai Tai HP; Newport Spectra Physics), a water-immersion objective ($\times 40$ LUMPlanFI/IR, 0.8 NA; Olympus), galvanometric scan mirrors (model 6210; Cambridge Technology), and a Pockel's Cell (Conoptics) for laser intensity modulation. For initial identification of *YC-Nano140*-expressing and tdTomato-positive S2P neurons, a volume stack was acquired using 800-nm excitation and yellow (542/50 nm) and red (610/75 nm) emission filters, respectively (AHF Analysentechnik). For calcium imaging, *YC-Nano140* was excited at 840 nm and fluorescence was collected with blue (480/60 nm) and yellow (542/50 nm) emission filters. Images were acquired at 7 Hz with 180×180 or 256×128 pixel resolution. Single trials of 6 to 7 s duration were recorded, with 1-s breaks between trials to allow for hard-disk storage during inter-trial periods.

Whisker tracking. The whisker field was illuminated with a 940-nm infrared LED light and videos were acquired at 500 Hz (500×500 pixels) using a high-speed CMOS camera (A504k; Basler). The average whisker angle across all imaged whiskers was measured using automated whisker-tracking software³⁵. For all trials, the duration of whisker to texture or pole contact was quantified manually through visual inspection. For a random subset of trials (approximately 80% of correct-rejection trials for texture discrimination, approximately 66% of hit and correct-rejection trials for object localization), the position of the principal whisker tip was tracked manually over a 1,100-ms time period starting 100 ms before the initial touch.

Post-hoc immunohistochemistry. After a period of five to seven days after CTB-Alexa647 injection (to allow uptake)³⁶, mice were anaesthetized (ketamine and xylazine; 100 mg per kg and 20 mg per kg body weight) and perfused transcardially with 4% paraformaldehyde in phosphate buffer, pH 7.4. Cortical sections (50 μ m) were cut along the imaging plane using a vibratome (VT100; Leica), then blocked in 10% normal goat serum (NGS) and 1% Triton at room temperature and incubated overnight at 4 °C in 5% NGS, 0.1% Triton, and the following primary antibodies: green fluorescent protein (GFP) (rat monoclonal antibody; 1:5,000; Nacalai) and GABA (rabbit polyclonal antibody; 1:1,000; Sigma). Appropriate Alexa488- and Alexa405-conjugated goat immunoglobulin-G (IgG) secondary antibodies (1:400; Molecular Probes, Invitrogen) were applied for 2 h at room temperature (approximately 22 °C). Images were acquired with a confocal microscope (Fluoview 1000; Olympus) with respective ultraviolet (GABA), green (*YC-Nano140*), red (tdTomato) and infrared (CTB-Alexa647) excitation or emission filters.

Calcium-imaging analysis. Two-channel (cyan fluorescent protein (CFP) and yellow fluorescent protein (YFP)) calcium-imaging data were imported into MATLAB (Mathworks) for processing. First, the background was subtracted on each channel (bottom 1st percentile fluorescence signal across the entire frame). Hidden Markov model line-by-line motion correction was applied to both data channels³⁷. Regions of interest (ROIs) corresponding to individual neurons were selected manually from the mean image of a single-trial time series using ImageJ (National Institutes of Health). Mean pixel value for each ROI was extracted for both channels. Calcium signals were expressed as relative YFP:CFP ratio change $\Delta R/R = (R - R_0)/R_0$. R_0 was calculated for each trial as the bottom 8th percentile of

the ratio for the trial. Active neurons were identified by two-way ANOVA with repeated measures of the neuronal calcium signal against the neuropil signal (significance value, $P < 0.05$). Owing to a number of factors—including variations in awake imaging conditions, variable noise levels across neuronal populations in the field of view, and slow decay kinetics of YC-Nano140 (Supplementary Fig. 3)—detection of calcium transients reflecting single action potentials is unlikely to occur with 100% fidelity during experimental conditions. Using a deconvolution approach, we obtained estimates of the action-potential firing-rate changes underlying the observed calcium signals as well as of the total number of extra action potentials evoked during the trial period (Supplementary Figs 3 and 8, and Supplementary Methods).

Behaviour classification. Behaviour vectors were constructed as follows. Owing to the relatively slow kinetic properties of YC-Nano140 and the low image acquisition rate (7 Hz) compared to observed whisking frequency (8 to 12 Hz)^{13,14,20,38}, measured calcium signals are unlikely to reflect whisking frequency accurately or to distinguish slow changes in whisking amplitude from fast rhythmic variation in position^{15,20}. As rhythmic and non-rhythmic whisking occurred (Supplementary Video 3), whisking amplitude was used as a measure to represent both forms of whisking behaviour. The envelope of whisking amplitude was calculated as the difference between maximum and minimum whisker angles along a sliding window equal to the imaging frame duration (142 ms). For whisker touch, a binary vector was constructed to represent whisker contact periods. Both vectors were downsampled to the 7-Hz imaging rate. Using MATLAB, cross correlation of behaviour vectors and cellular calcium signal vectors was performed across the entire behaviour session, with all trials concatenated into single vectors with n elements. For two time-series vectors, \mathbf{x} and \mathbf{y} , covariance across lags $k = 0, \pm 1, \pm 2$ frames was calculated as:

$$c_{xy} = \begin{cases} \frac{1}{n} \sum_{t=1}^{n-k} (x_t - \bar{x})(y_{t+k} - \bar{y}) & k = 0, 1, 2 \dots \\ \frac{1}{n} \sum_{t=1}^{n+k} (y_t - \bar{y})(x_{t-k} - \bar{x}) & k = -1, -2 \dots \end{cases}$$

where \bar{x} and \bar{y} denote the means of \mathbf{x} and \mathbf{y} across all values. The sample cross-correlation (R value) was then computed as:

$$R_{xy}(k) = \frac{c_{xy}(k)}{\sigma_x \sigma_y} \quad k = 0, \pm 1, \pm 2 \dots$$

where σ_x and σ_y denote standard deviations of \mathbf{x} and \mathbf{y} , respectively. To determine whether calcium signals of a neuron were significantly correlated to a given behaviour above chance levels, we carried out a random permutation test by shuffling individual time points along the calcium signal vector across the entire behaviour session. The sample cross correlation was then computed on this shuffled data set, and shuffling was repeated 1,000 times to obtain a distribution of R values from which confidence intervals and P values could be determined for significance tests.

To distinguish putative 'touch' and 'whisking' cells, we compared the R values from touch and whisking vectors statistically by bootstrapping with sample replacement. Calcium transients from individual trials were replaced randomly with transients of other trials from the same behaviour session. The sample cross correlation was then computed against the touch and whisking vector on this bootstrapped data set. This process was repeated 1,000 times to obtain 95% confidence intervals for significance tests. In general, correlations with positive time lags indicate that calcium signals follow behaviour, whereas correlations with negative time lags indicate that calcium signals precede behaviour (specifically, whisking cells are typically active before touch; Supplementary Fig. 8). The peak R value over the range of time lags analysed typically correlates to the peak of the calcium transient, as opposed to the calcium signal onset. Touch cells were identified as having a peak positive R value for touch above chance (from random permutation test) between lags $k = 0$ to 4 frames ($t = 0$ to 571 ms). This R value had to be significantly higher than the R value for whisking (from bootstrap test) at the corresponding time lag. As whisking signals show a delayed time to peak correlation, compared to touch signals¹⁸, whisking cells were identified as having a peak positive R value for whisking between lags $k = 2$ to 5 frames ($t = 285$ to 714 ms). This R value had to be significantly higher than the R value for touch at the corresponding time lag. Cells with no significant correlation for touch or whisking were unclassified. Cells with a peak negative R value between lags $k = 0$ to 5 frames or a peak positive R value at $k > 5$ frames or $k < 0$ frames for both touch and whisking were also unclassified.

To determine the likelihood that the measured distribution of behaviour-classified neurons across S2P, M1P and unlabelled subtypes was above chance, a random permutation test was performed by shuffling the classification labels. For each permutation, the fraction of cells belonging to a particular behaviour classification was calculated. This process was repeated 1,000 times to obtain a distribution for each behaviour-subtype combination representing the null hypothesis that no

behaviour-subtype dependencies exist. To compare the measured distribution of classified cells across S2P, M1P and unlabelled subtypes, bootstrapping with sample replacement of cells with a defined behaviour-subtype combination was performed 1,000 times to obtain 95% confidence intervals for each observed combination. The bias index for individual neurons in naive animals was calculated as $(R_{\text{sandpaper}} - R_{\text{pole}})/(R_{\text{sandpaper}} + R_{\text{pole}})$, where the R value represents the peak correlation within lags $k = 0$ to 4 frames ($t = 0$ to 571 ms) of the calcium signal to touch vector.

Whisker kinematic analysis. For analysis of kinematic parameters, 'slip' events were identified as described previously¹¹. Principal whisker velocity and acceleration were determined by calculating the first and second derivative, respectively, of the time vector representing whisker-tip displacement. Slips were defined as events with both high positive acceleration and high absolute velocity, exceeding thresholds corresponding to four standard deviations above the means across all measured trials ($\theta_A = 20.9 \text{ mm ms}^{-2}$ and $\theta_V = 35.6 \text{ mm ms}^{-1}$). For texture discrimination, slips represent motion occurring as a result of the contact between whisker tip and texture surface as described previously¹¹. During object localization, pole contact typically occurs along the whisker shaft. In this case, a high-acceleration, high-velocity event satisfying the criteria of a slip event occurs either following initial pole contact, producing a fast recoiling motion, or during full whisking behaviour when the whisker initially catches but bends past the pole, springing free. Whisker curvature change ($|\Delta\kappa|$), reflecting contact forces upon touch³⁹, was measured as the whisker curvature at each time point of touch minus the mean curvature over the 100-ms period before initial touch. Whisker curvature was determined as $\kappa = 1/R$, where R is the radius of a circle, in which the whisker length represents the arc length between the follicle and tip position and in which the direct distance between the follicle and tip position represents the chord. Whisker angle was calculated as the arctangent of $(x_{\text{follicle}} - x_{\text{tip}})/(y_{\text{follicle}} - y_{\text{tip}})$, where 0 degrees is orthogonal to the anterior-posterior axis.

For cross-correlation analysis of kinematic parameters to calcium signals, time vectors of kinematic parameters were downsampled to the imaging frame rate and analysis was carried out on the first second after initial touch (7 frames per trial) for correct-rejection trials. For slips, cross correlation was performed on a time vector corresponding to the cumulative number of slips over the time period. For curvature change, cross correlation was performed on a time vector representing the max $|\Delta\kappa|$ at each frame interval. For whisker angle, the time vector representing whisker angle was separated into eight binary vectors representing a range of angles from -90° to 90° , subdivided at 22.5° intervals. Each sub-vector reflects the likelihood that the whisker angle falls within that given angle range. Cross correlation was performed on each sub-vector and the overall correlation to whisker angle was obtained by taking the maximum R value across the eight sub-vectors. Owing to limited imaging speed, slow calcium-indicator kinetics, and because L2/3 neuronal activity in S1 is largely devoid of fast touch signals^{15,16}, measured calcium responses are not likely to represent instantaneous kinematic features. The correlation values obtained represent relative measures of response preferences to these parameters solely to permit a comparison across cell types.

Single-neuron discrimination analysis. The performance of single neurons in discriminating behavioural decisions or presented stimuli was assessed using a receiver operating characteristic (ROC) analysis^{16,21}. Classification of decision X (or stimulus X) versus decision Y (stimulus Y) was based on the similarity of the calcium transient in each trial to the mean calcium transient for trial type X compared to trial type Y . Only the first second of the calcium signals following initial texture or pole contact was considered. Each trial was assigned a 'discrimination variable' score (DV) equal to the dot-product similarity to the mean calcium transient for trial type X minus the dot-product similarity to the mean for trial type Y . Thus, for trial type X :

$$DV_X = X_i (\bar{X}_{vj \neq i} - \bar{Y})$$

and for trial type Y :

$$DV_Y = Y_i (\bar{X} - \bar{Y}_{vj \neq i})$$

where X_i and Y_i are the single-trial calcium transients for the i -th trial. \bar{X} and \bar{Y} are the mean 1-s calcium transients after initial texture or pole contact for the respective trial type (with the trial under consideration omitted for averaging). Trials were classified as belonging to trial type X or Y if DV_X or DV_Y was greater than a given criterion, respectively. To determine the fraction of trials an ideal observer could correctly classify, a ROC curve was constructed by varying this criterion value across the entire range of DV_X or DV_Y . At each criterion value, the probability that a trial of type X exceeded the criterion value was plotted against the probability that a trial of type Y exceeded the criterion value. The area under the ROC curve was then calculated to represent the single-neuron performance ('fraction correct') as the fraction of trials correctly discriminated by an ideal observer using the DV ²¹.

Neurons that discriminated above chance were identified using repeated permutations tests in which decision or stimulus labels were randomly shuffled. For each permutation test, a threshold corresponding to the shuffled distribution 0.95 percentile was calculated. Neurons, whose performance values were above the mean value of this threshold across 1000 permutation tests, were considered to be discriminating above chance.

31. Gradinaru, V. *et al.* Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* **141**, 154–165 (2010).
32. Lütcke, H. *et al.* Optical recording of neuronal activity with a genetically-encoded calcium indicator in anesthetized and freely moving mice. *Front Neural Circuits* **4**, 9 (2010).
33. Margolis, D. J. *et al.* Reorganization of cortical population activity imaged throughout long-term sensory deprivation. *Nature Neurosci.* **15**, 1539–1546 (2012).
34. Langer, D. *et al.* HelioScan: A software framework for controlling in vivo microscopy setups with high hardware flexibility, functional diversity and extendibility. *J. Neurosci. Methods* **215**, 38–52 (2013).
35. Knutsen, P. M., Derdikman, D. & Ahissar, E. Tracking whisker and head movements in unrestrained behaving rodents. *J. Neurophysiol.* **93**, 2294–2301 (2004).
36. Conte, W. L., Kamishina, H. & Reep, R. L. Multiple neuroanatomical tract-tracing using fluorescent Alexa Fluor conjugates of cholera toxin subunit B in rats. *Nature Protocols* **4**, 1157–1166 (2009).
37. Dombeck, D. A., Khabbazi, A. N., Collman, F., Adelman, T. L. & Tank, D. W. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. *Neuron* **56**, 43–57 (2007).
38. de Kock, C. P. & Sakmann, B. Spiking in primary somatosensory cortex during natural whisking in awake head-restrained rats is cell-type specific. *Proc. Natl Acad. Sci. USA* **106**, 16446–16450 (2009).
39. Birdwell, J. A. *et al.* Biomechanical models for radial distance determination by the rat vibrissa system. *J. Neurophysiol.* **98**, 2439–2455 (2007).

Ratios of S, Se and Te in the silicate Earth require a volatile-rich late veneer

Zaicong Wang¹ & Harry Becker¹

The excess of highly siderophile (iron-loving) elements (HSEs) and the chondritic ratios of most HSEs in the bulk silicate Earth (BSE) may reflect the accretion of a chondritic 'late veneer' of about 0.5 per cent of Earth's mass after core formation^{1,2}. The amount of volatiles contained in the late veneer is a key constraint on the budget and the origin of the volatiles in Earth. At high pressures and temperatures, the moderately volatile chalcogen elements sulphur (S), selenium (Se) and tellurium (Te) are moderately to highly siderophile; thus, if depleted by core formation their mantle abundances should reflect the volatile composition of the late veneer^{3,4}. Here we report ratios and abundances of S, Se and Te in the mantle determined from new isotope dilution data for post-Archaean mantle peridotites. The mean S/Se and Se/Te ratios of mantle lherzolites overlap with CI (Ivuna-type) carbonaceous chondrite values^{5,6}. The Se/Te ratios of ordinary and enstatite chondrites are significantly different. The chalcogen/HSE ratio of the BSE is similar to that of CM (Mighei-type) carbonaceous chondrites, consistent with the view that the HSE signature of the BSE reflects a predominance of slightly volatile-depleted, carbonaceous-chondrite-like material, possibly with a minor proportion of non-chondritic material⁷. Depending on the estimates for the abundances of water and carbon in the BSE⁸, the late veneer may have supplied 20 to 100 per cent of the budget of hydrogen and carbon in the BSE.

Understanding the history of the delivery to Earth of volatile elements, particularly hydrogen and carbon, is of fundamental importance for understanding planetary evolution⁹. Earth formed during collisions of protoplanetary embryos with compositions that may have evolved from being reducing and poor in volatiles to being oxidizing and presumably rich in volatiles^{10–12} (the heterogeneous accretion model). The detailed origin and history of moderately volatile and atmophile elements in Earth remains uncertain, mostly as a result of considerable element redistribution during core formation and silicate differentiation^{8,11,13,14}. It is likely that a fraction of the atmophile-element budget of the Earth was delivered after core formation, but little agreement exists on the sources and mass fractions of volatiles delivered late^{13,14}. Some information about the composition of materials that accreted after core formation (forming the 'late veneer') has been derived from highly siderophile elements^{1,7,15,16} (the platinum group elements, Re and Au). During core formation the HSEs should have partitioned into the metallic core of Earth nearly quantitatively, leaving the BSE strongly depleted in HSEs, which is not the case. Experiments at high pressures and temperatures indicate significantly lower partition coefficients than previously thought for the HSEs during terrestrial core formation, but a late veneer is still required².

With the exception of Pd and Au, which are slightly volatile, most HSEs are refractory elements. The ¹⁸⁷Os/¹⁸⁸Os ratio of the BSE reflects its time-integrated Re/Os ratio, which has been interpreted such that the late veneer may have been mostly comprised of water- and carbon-depleted, ordinary or enstatite chondrites^{17,18}. The BSE is characterized by excesses of Pd and Ru of 40–50% and 30%, respectively, relative to Ir, Os and Pt (ref. 15). The Pd excess may be attributed to mixing of a chondritic late veneer with mantle that has a small Pd excess inherited

from core–mantle equilibrium partitioning; however, the excess of Ru can be explained neither by metal–silicate segregation at high pressures and temperatures nor by a chondritic late veneer^{1,2,15}. These data imply that a chondritic late veneer alone cannot explain the composition of the BSE: either additional processes are required or the late veneer may have had a somewhat different refractory-element composition compared with known chondrites^{1,15}. HSE data from >3.8-Gyr-old lunar impact rocks related to large basin-forming impacts provide hints that the late veneer accreted to the Earth may have been comprised of carbonaceous-chondrite-like material with a minor contribution from a non-chondritic component⁷. These data suggest that if the late veneer represents a mixture of materials with different compositions, the ¹⁸⁷Os/¹⁸⁸Os ratio and HSE composition of the BSE cannot provide stringent constraints on the composition and class of the chondritic material in the late veneer.

Sulphur, selenium and tellurium have similar, and quite low, 50% condensation temperatures near 700 K in a canonical solar nebula composition⁶. An important property of the chalcogens is that they become more siderophile with increasing pressure and are moderately to highly siderophile during metal–silicate segregation at high pressures and temperatures⁴. As for the HSEs, the experimental data predict that S, Se and Te in the BSE should predominantly derive from the late veneer, if metal–silicate segregation occurred at high pressure and temperature and under equilibrium conditions⁴. Because different classes of chondrites display the variable fractionation behaviour of moderately volatile elements, the chalcogens can distinguish contributions from carbonaceous-, ordinary- or enstatite-chondrite-like materials in the late veneer. Notably, different carbonaceous-chondrite groups all have Se/Te ratios similar to the chondrites that are the chemically most primitive (CI chondrites, Se/Te = 8.5 ± 0.7 (all uncertainties are s.d. unless specified otherwise); ref. 6), irrespective of the extent of depletion of the chalcogens; whereas ordinary and enstatite chondrites are characterized by significantly higher Se/Te ratios (11–30) than are carbonaceous chondrites (Supplementary Table 6). In addition, the abundance of the chalcogens relative to the refractory HSEs in the BSE may constrain the depletion of the chalcogens in the late veneer, the type of primitive material and its likely contribution to highly volatile elements such as hydrogen and carbon.

Abundances of the chalcogens in the BSE have been estimated previously on the basis of a limited number of mantle peridotites. However, owing to analytical difficulties, abundance estimates of Se and Te in the BSE were based on very little data, and reliable data for S/Se and Se/Te ratios have therefore been scarce^{19–22}. To obtain an improved database for the chalcogen composition of the mantle, we present new, precise S, Se and Te concentration data obtained by isotope dilution for the same digestion aliquot of fresh mantle peridotites (Figs 1 and 2 and Supplementary Table 2). The new data will be compared with lithophile-element data and precise HSE data obtained from the same samples, to evaluate the behaviour of the chalcogens relative to these elements during mantle processes.

As in previous work^{15,19,22,23}, the data show a linear decrease in sulphur abundance with decreasing Al₂O₃ content, an indicator of

¹Freie Universität Berlin, Institut für Geologische Wissenschaften, Malteserstrasse 74-100, 12249 Berlin, Germany.

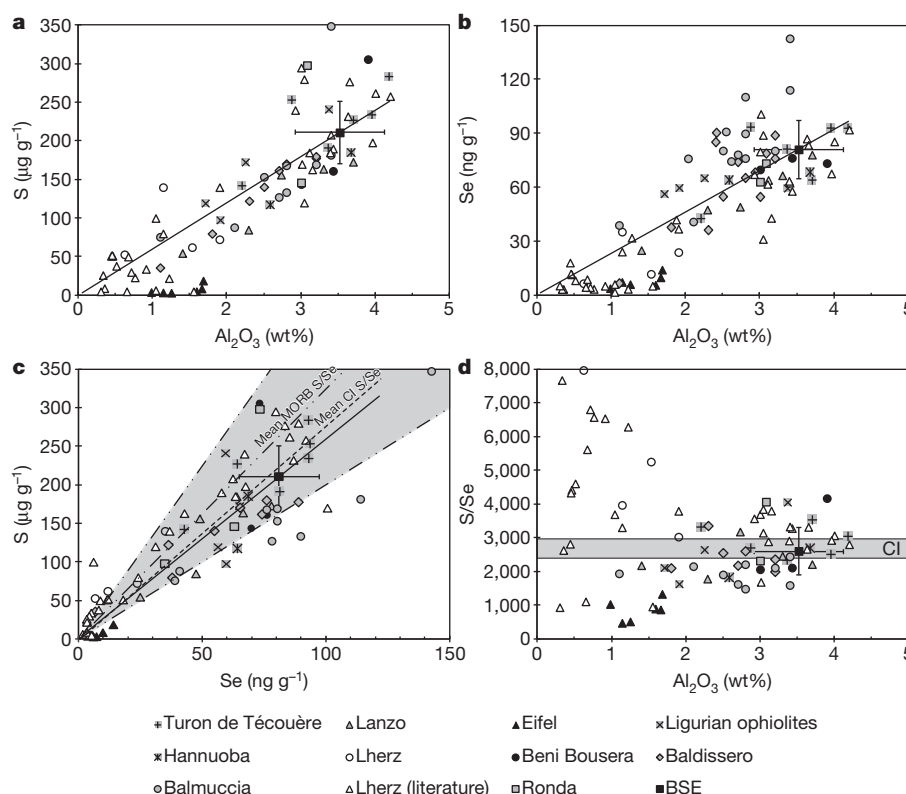


Figure 1 | Correlation of S and Se contents of terrestrial peridotites with Al_2O_3 contents as indicator of depletion. a–c, Sulphur, Se and Al_2O_3 are positively correlated with each other for mantle peridotites from different geological settings (for locations see legend). Lines in a–c are linear regression lines of the data. Shaded region in c is the range of S/Se ratios of mid-ocean-ridge basalts (MORBs), with a mean value of $3,230 \pm 560$ ($n = 279$; ref. 28), which is somewhat higher than the mean S/Se ratio of mantle peridotites. d, S/Se ratios in lherzolites ($\text{Al}_2\text{O}_3 > 1.5$ wt%) show scatter, but change little with depletion. In contrast, S/Se ratios in harzburgites are much more variable.

The mean S/Se ratio ($2,690 \pm 700$, $n = 53$) of the lherzolites is indistinguishable from the mean CI chondrite value ($2,750 \pm 200$; ref. 6). The Al_2O_3 concentration of the BSE given here is based on recent statistical evaluation of peridotite data²⁵, and should be viewed as a minimum value (Supplementary Information). Literature data for the Lherz massif are from refs 22, 29, and data for the Balmuccia and Baldissero massifs are from ref. 23. Precisions of chalcogen concentrations are within symbol size and error bars of the BSE values are 1 s.d.

progressive depletion by melting, both for the whole data set of mantle peridotites and for individual suites of peridotites (Fig. 1a). Notably, the linear correlation of S with Al_2O_3 occurs in suites of peridotites that display different states of equilibration of $^{187}\text{Os}/^{188}\text{Os}$ ratios and HSE abundances^{16,22,23}. Similarly, Se and Te abundances are also positively correlated with Al_2O_3 and S contents, although with larger scatter (Figs 1 and 2). The correlation between the chalcogens and Al_2O_3 and incompatible HSEs (Supplementary Information) and the limited fractionation of S/Se and Se/Te in lherzolites provide a solid foundation for the determination of ratios and abundances of these elements in the BSE. Sulphur/selenium ratios in lherzolites ($\text{Al}_2\text{O}_3 > 1.5$ wt%) display some scatter with a mean of $S/\text{Se} = 2,690 \pm 700$ ($n = 53$, including the Lherz massif²²), consistent with a value of $2,600 \pm 700$ derived from linear regression of the correlation of S with Se (Fig. 1c and Supplementary Table 5). Selenium/tellurium ratios yield a mean of $\text{Se}/\text{Te} = 7.9 \pm 1.6$ ($n = 63$) and slightly decrease with depletion, implying that Se is slightly more incompatible than Te (Fig. 2c).

The scatter in chalcogen-element abundances and ratios between different suites of fertile lherzolites probably reflects different extents of magmatic re-equilibration of mantle domains during melt migration and open-system partial melting²³, and, thus, mantle heterogeneity. Infiltration and trapping of silicate melt in peridotitic melting residues are now recognized as widespread processes in the mantle²⁴. Petrographic, major-element, HSE and $^{187}\text{Os}/^{188}\text{Os}$ data indicate that many samples of this study probably were affected by such processes^{16,23}. Nonetheless, the effects of these processes on abundances of S, Se and Te seem to be very similar to those of equilibrium partitioning

during open-system melting²³. The limited variations suggest that the average S/Se and Se/Te ratios of lherzolites should provide robust constraints on the composition of the BSE, irrespective of the detailed mantle processes.

Mantle heterogeneity only moderately affects linear correlations of S and Se with Al_2O_3 in peridotites (Fig. 1). Thus, regressions can constrain abundances of S and Se in the BSE, provided that the Al_2O_3 content of the BSE is known. A recent re-evaluation of the Al_2O_3 content of the BSE yielded a value of 3.52 ± 0.60 wt% (ref. 25), which is 20–30% lower than previous estimates (4.0–4.5 wt% Al_2O_3 ; refs 20, 21). This lower Al_2O_3 content yields minimum estimates for S and Se abundances in the BSE. The uncertainty in Al_2O_3 has negligible effects on the estimate of S/Se in the BSE (Fig. 1 and Supplementary Table 5). After excluding data from mantle xenoliths that have been affected by secondary processes (Supplementary Information), the lower value for Al_2O_3 in the BSE yields minimum S and Se abundances of the BSE that are slightly lower, but still within the uncertainty of previous estimates ($S = 211 \pm 40 \mu\text{g g}^{-1}$ and $\text{Se} = 80 \pm 17 \text{ ng g}^{-1}$, or 0.0040 ± 0.0008 times the values in CI chondrites; refs 20, 21).

Tellurium is more compatible than Al during igneous processes in the mantle²³, and abundances thus display more scatter for a given Al_2O_3 content (Supplementary Fig. 1). Tellurium/palladium ratios in mantle peridotites decrease with decreasing Te contents, indicating that Te is more incompatible than Pd during mantle processes (Fig. 2d). The relative sequence of empirical bulk partition coefficients of residue/melt during magmatic processes in the upper mantle is thus $D_{\text{Pd}} > D_{\text{Te}} > D_{\text{Se}} > D_{\text{S}} \approx D_{\text{Re}} \approx D_{\text{Al}}$ (ref. 23; Figs 1 and 2 and

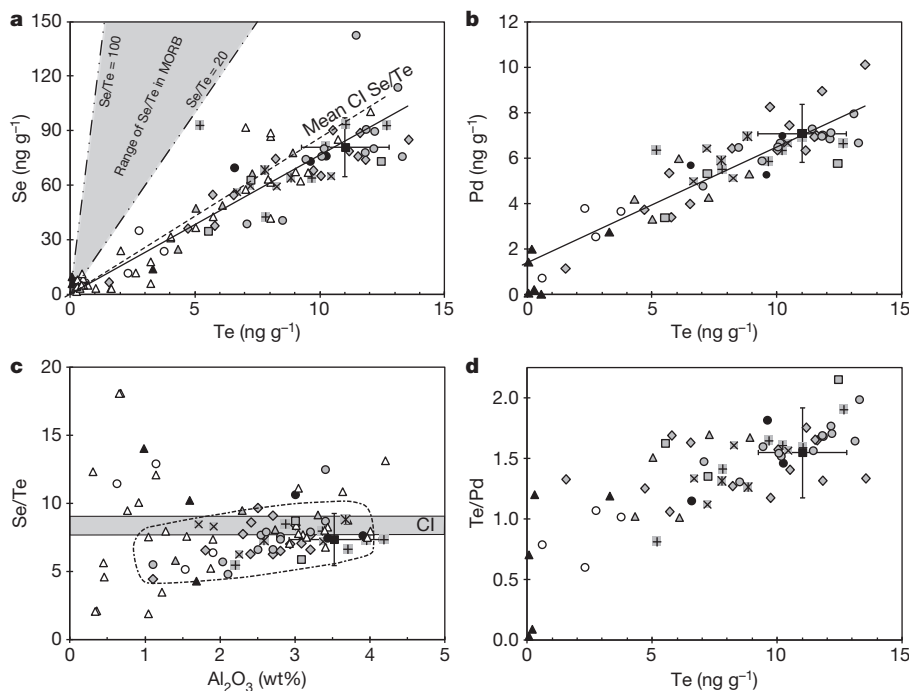


Figure 2 | Correlations of Te with Se and Pd contents. **a**, **b**, Tellurium shows positive correlations with Se and Pd (regression lines are solid). Dashed lines in **a** labelled Se/Te = 20 and 100 reflect the range of Se/Te ratios in MORBs for comparison³⁰. **c**, In lherzolites, the Se/Te ratio decreases slightly with decreasing Al₂O₃ (outlined region) but the mean value of fertile (Al₂O₃-rich) lherzolites overlaps with CI chondrite data within the uncertainties (mean CI Se/Te = 8.5 ± 0.7 ; ref. 6). Only a few lherzolites have higher Se/Te ratios than the CI value. As they do for S/Se, harzburgites have highly variable Se/Te ratios at low concentrations of these elements. **d**, Te/Pd ratios decrease with decreasing Te content, indicating that Te is more incompatible than Pd. Symbol key and minimum Al₂O₃ content of BSE as in Fig. 1. Precisions of concentrations are similar or smaller than symbol size and error bars of the BSE are 1 s.d.

Supplementary Fig. 2). This sequence also seems to be valid for peridotites affected by melt infiltration. Because the incompatibility of Te lies between those of Pd and Se, upper and lower bounds on Te abundances in the BSE were obtained from regression of correlations with these elements ($\text{Te} = 11.0 \pm 1.7 \text{ ng g}^{-1}$; Fig. 2 and Supplementary Tables 4 and 5).

High metal–silicate partition coefficients during core formation at high pressures and temperatures indicate that the chalcogens would

have been almost completely lost to the core⁴. If significant amounts of chalcogen had been retained in the BSE after core formation, they would display non-chondritic ratios⁴; the carbonaceous-chondrite-like ratios of the chalcogens in the BSE therefore support the late veneer as their main source. Within their uncertainties, the S/Se and Se/Te ratios of the BSE agree with values of CI chondrites (Figs 1–3 and Supplementary Tables 5 and 6), and accordingly, the late veneer must have consisted mostly of carbonaceous-chondrite-like material. The data in Fig. 2 hint that the Se/Te ratio of the BSE may be 10% lower than that of CI chondrites, but a difference cannot be resolved given the scatter of the data. CI-chondrite-normalized S/Ir, Se/Ir and Te/Ir ratios are similar to each other and match the volatile-element depletion trend of CM chondrites but not that of H (high-iron-type) chondrites or any other ordinary chondrite group²¹ (Fig. 3).

The carbonaceous-chondrite-like chalcogen composition of the BSE lends further support to recent models of the composition of the late veneer based on HSE ratios in the BSE and in ancient lunar impact rocks. A mixed provenance of the late veneer with a large fraction of carbonaceous-chondrite-like material and a minor proportion of non-chondritic material—possibly differentiated meteoritic metal—could be a likely explanation⁷. The minor contribution from metal that might explain the HSE systematics in the BSE may have limited effects on the chalcogens (Supplementary Information). Alternatively, relatively volatile-rich, primitive material with suprachondritic Ru/Ir and Pd/Ir ratios may explain the HSE composition of the BSE^{7,15}. Material with such a composition is different from known primitive meteorites.

The CM-chondrite-like chalcogen composition of the late veneer yields constraints on the origin and timing of delivery of highly volatile (atmophile) elements to Earth. If all sulphur of the late veneer had been

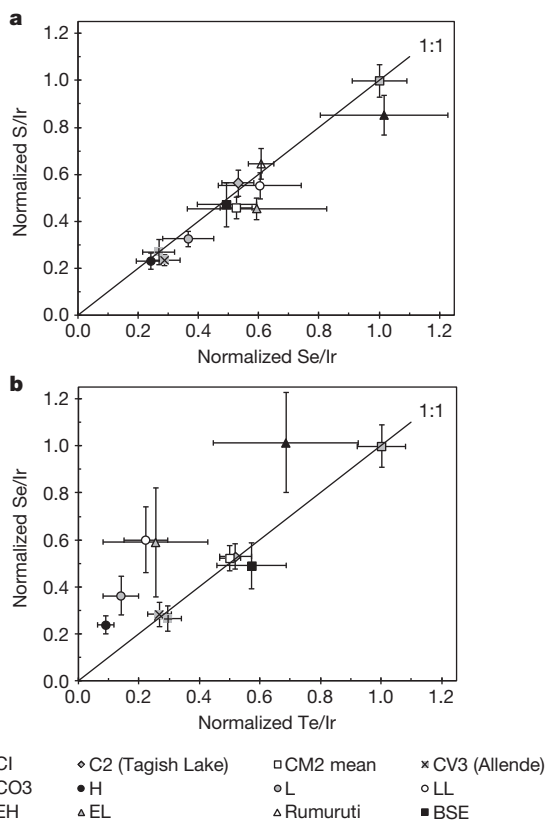


Figure 3 | CI-chondrite-normalized ratios of moderately volatile chalcogens relative to the refractory highly siderophile element Ir. Data are shown for different groups of chondrites and the BSE (Supplementary Table 6). **a**, **b**, Because of the depletion of the chalcogens relative to the HSEs, the BSE has S/Ir, Se/Ir and Te/Ir ratios similar to CM chondrites. Ordinary and enstatite chondrites lie significantly above the 1:1 line in **b**, because they have significantly higher Se/Te ratios or they have different chalcogen/HSE ratios than do carbonaceous chondrites and the BSE. Error bars, 1 s.d. The 1:1 lines indicate CI-chondrite-like chalcogen–element ratios.

delivered by comets, the amounts of atmophile elements added would exceed most estimates of the abundances of such elements in the BSE and should have resulted in significantly lower H/C ratios in the BSE (Supplementary Table 7). The D/H and $^{15}\text{N}/^{14}\text{N}$ ratios of Earth's surface inventory are similar to CI and CM chondrites, but different from solar and presumably also cometary compositions^{8,26}. However, considerable uncertainties exist regarding the ranges of light, stable isotope ratios in comets⁸. Combined isotopic and compositional data indicate that carbonaceous-chondrite-like material was the most likely source of a large fraction of Earth's noble gases, hydrogen, nitrogen and presumably also carbon^{8,26}.

Conservative estimates of the abundances of water and carbon in the mantle and in surface reservoirs imply that the late veneer may be the dominant source of water and carbon in the BSE (Supplementary Information). By comparison with noble-gas budgets, it has been suggested that the abundances of water and carbon in the BSE may be as high as 1,000–3,000 $\mu\text{g g}^{-1}$ and 500 $\mu\text{g g}^{-1}$, respectively, equivalent to $2 \pm 1\%$ of Earth's mass of CI- or CM-chondrite-like material⁸. In this case, the contribution of the late veneer to Earth's budget of water and carbon may be only 20–50%. If correct, these estimates require accretion of CI- or CM-chondrite-like materials during core formation. Larger quantities of atmophile elements in Earth's interior would be consistent with the proposed accretion of volatile-bearing carbonaceous-chondrite-like material during the final stages of core formation, as suggested by modelling of Ag isotopic data¹¹ and heterogeneous accretion models¹². A late veneer predominantly composed of volatile-rich carbonaceous-chondrite-like materials may then simply reflect the tail of the late phase of terrestrial planet accretion, as proposed in recent work on dynamical models²⁷.

METHODS SUMMARY

Concentrations of S, Se, Te and HSEs in bulk rocks of peridotites and some carbonaceous chondrites (Supplementary Information) were determined by isotope dilution and inductively coupled plasma mass spectrometry (ICP-MS) using high-pressure asher (HPA-S) digestion. The details have been described elsewhere^{16,23}. About 2.5 g of peridotite powder or 50–70 mg of meteorite sample powder were digested by HPA-S in 5 ml of 14 M HNO_3 and 2.5 ml of 9 M HCl at 100 bar and 320 °C for 16 h after addition of mixed ^{77}Se – ^{125}Te and ^{34}S spike solutions for all samples, and a mixed ^{191}Ir – ^{99}Ru – ^{194}Pt – ^{105}Pd spike solution for some samples to obtain duplicate data for comparison with literature values. After digestion, one third of the digestion solution was used for chemical separation of Ir, Ru, Pt and Pd as described elsewhere¹⁶, and about one-quarter of the remaining solution was used for S–Se–Te separation.

The S–Se and Te fractions were obtained by separation on anion-exchange resin using 6 M HCl, 9 M HCl, 5 M hydrofluoric acid and 1 M HNO_3 (ref. 23). The separated S–Se fraction was purified on cation-exchange resin using 0.1 M HNO_3 . Sulphur measurements were performed on the S–Se fraction using a Scott-type glass spray chamber in medium-mass-resolution mode ($m/\Delta m = 4,000$) on the Element XR ICP-MS at Freie Universität Berlin²³. Selenium and Te were measured in low-mass-resolution mode on the Element XR, using a double-pass Scott-type glass spray chamber, combined with a hydride-generation sample-introduction system by reacting the sample solution with 1% m/m NaBH_4 in 0.05 M NaOH (ref. 23). Procedural blanks for the chalcogens were in most cases insignificant ($<1\%$; $3 \pm 1.5 \mu\text{g S}$, $1.5 \pm 1 \text{ ng Se}$ and $20 \pm 15 \text{ pg Te}$).

Full Methods and any associated references are available in the online version of the paper.

Received 10 December 2012; accepted 8 May 2013.

1. Walker, R. J. Highly siderophile elements in the Earth, Moon and Mars: update and implications for planetary accretion and differentiation. *Chem. Erde Geochem.* **69**, 101–125 (2009).
2. Mann, U., Frost, D. J., Rubie, D. C., Becker, H. & Audetat, A. Partitioning of Ru, Rh, Pd, Re, Ir and Pt between liquid metal and silicate at high pressures and high temperatures: implications for the origin of highly siderophile element concentrations in the Earth's mantle. *Geochim. Cosmochim. Acta* **84**, 593–613 (2012).
3. Yi, W. *et al.* Cadmium, indium, tin, tellurium, and sulfur in oceanic basalts: Implications for chalcophile element fractionation in the Earth. *J. Geophys. Res. Solid Earth* **105**, 18927–18948 (2000).

4. Rose-Weston, L., Brennan, J. M., Fei, Y. W., Secco, R. A. & Frost, D. J. Effect of pressure, temperature, and oxygen fugacity on the metal-silicate partitioning of Te, Se, and S: implications for earth differentiation. *Geochim. Cosmochim. Acta* **73**, 4598–4615 (2009).
5. Dreibus, G., Palme, H., Spettel, B., Zipfel, J. & Wänke, H. Sulfur and selenium in chondritic meteorites. *Meteoritics* **30**, 439–445 (1995).
6. Lodders, K. Solar system abundances and condensation temperatures of the elements. *Astrophys. J.* **591**, 1220–1247 (2003).
7. Fischer-Gödde, M. & Becker, H. Osmium isotope and highly siderophile element constraints on ages and nature of meteoritic components in ancient lunar impact rocks. *Geochim. Cosmochim. Acta* **77**, 135–156 (2012).
8. Marty, B. The origins and concentrations of water, carbon, nitrogen and noble gases on Earth. *Earth Planet. Sci. Lett.* **313–314**, 56–66 (2012).
9. Abe, Y., Ohtani, E., Okuchi, T., Righter, K. & Drake, M. in *Origin of the Earth and Moon* (eds R. M. Canup & K. Righter) 413–433 (Univ. Arizona Press, 2000).
10. Wood, B. J., Walter, M. J. & Wade, J. Accretion of the Earth and segregation of its core. *Nature* **441**, 825–833 (2006).
11. Schönbachler, M., Carlson, R. W., Horan, M. F., Mock, T. D. & Hauri, E. H. Heterogeneous accretion and the moderately volatile element budget of Earth. *Science* **328**, 884–887 (2010).
12. Rubie, D. C. *et al.* Heterogeneous accretion, composition and core-mantle differentiation of the Earth. *Earth Planet. Sci. Lett.* **301**, 31–42 (2011).
13. Albarède, F. Volatile accretion history of the terrestrial planets and dynamic implications. *Nature* **461**, 1227–1233 (2009).
14. Wood, B. J. & Halliday, A. N. The lead isotopic age of the Earth can be explained by core formation alone. *Nature* **465**, 767–770 (2010).
15. Becker, H. *et al.* Highly siderophile element composition of the Earth's primitive upper mantle: constraints from new data on peridotite massifs and xenoliths. *Geochim. Cosmochim. Acta* **70**, 4528–4550 (2006).
16. Fischer-Gödde, M., Becker, H. & Wombacher, F. Rhodium, gold and other highly siderophile elements in orogenic peridotites and peridotite xenoliths. *Chem. Geol.* **280**, 365–383 (2011).
17. Walker, R. J. *et al.* Comparative ^{187}Re – ^{187}Os systematics of chondrites: implications regarding early solar system processes. *Geochim. Cosmochim. Acta* **66**, 4187–4201 (2002).
18. Meisel, T., Walker, R. J., Irving, A. J. & Lorand, J. P. Osmium isotopic compositions of mantle xenoliths: a global perspective. *Geochim. Cosmochim. Acta* **65**, 1311–1323 (2001).
19. Morgan, J. W. Ultramafic xenoliths: clues to Earth's late accretionary history. *J. Geophys. Res. Solid Earth* **91**, 12375–12387 (1986).
20. McDonough, W. F. & Sun, S. S. The composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).
21. Palme, H. & O'Neill, H. S. C. in *Treatise on Geochemistry* Vol. 2 (eds H. D. Holland & K. K. Turekian) 1–38 (Pergamon, 2003).
22. Lorand, J. P. & Alard, O. Determination of selenium and tellurium concentrations in Pyrenean peridotites (Ariege, France): new insight into S/Se/Te systematics of the upper in mantle samples. *Chem. Geol.* **278**, 120–130 (2010).
23. Wang, Z., Becker, H. & Gawronski, T. Partial re-equilibration of highly siderophile elements and the chalcogens in the mantle: a case study on the Baldissero and Balmuccia peridotite massifs (Ivrea Zone, Italian Alps). *Geochim. Cosmochim. Acta* **108**, 21–44 (2013).
24. Bodinier, J. L. & Godard, M. in *Treatise on Geochemistry* Vol. 2 (eds H. D. Holland & K. K. Turekian) 103–170 (Pergamon, 2003).
25. Lyubetskaya, T. & Korenaga, J. Chemical composition of Earth's primitive mantle and its variance: 1. Method and results. *J. Geophys. Res. Solid Earth* **112**, B03211 (2007).
26. Alexander, C. M. O. D. *et al.* The provenances of asteroids, and their contributions to the volatile inventories of the terrestrial planets. *Science* **337**, 721–723 (2012).
27. Bottke, W. F., Walker, R. J., Day, J. M. D., Nesvorniy, D. & Elkins-Tanton, L. Stochastic late accretion to Earth, the Moon, and Mars. *Science* **330**, 1527–1530 (2010).
28. Jenner, F. E. & O'Neill, H. S. C. Analysis of 60 elements in 616 ocean floor basaltic glasses. *Geochim. Geophys. Geosyst.* **13**, Q02005 (2012).
29. König, S., Luguet, A., Lorand, J.-P., Wombacher, F. & Lissner, M. Selenium and tellurium systematics of the Earth's mantle from high precision analyses of ultra-depleted orogenic peridotites. *Geochim. Cosmochim. Acta* **86**, 354–366 (2012).
30. Hertogen, J., Janssens, M. J. & Palme, H. Trace elements in ocean ridge basalt glasses: implications for fractionations during mantle evolution and petrogenesis. *Geochim. Cosmochim. Acta* **44**, 2125–2143 (1980).

Supplementary Information is available in the online version of the paper.

Acknowledgements Some peridotite and meteorite samples were provided by S. Gao, J.-P. Lorand, G. MacPherson and M. Wadhwa. We thank F. Wombacher and C. Funk for discussions and M. Feth, K. Hammerschmidt and M. Weynell for technical assistance. This work was supported by funds from Freie Universität Berlin and a China Scholarship Council fellowship to Z.W.

Author Contributions Z.W. and H.B. wrote the paper. H.B. designed the project. Z.W. developed the analytical methods and performed the analyses.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Z.W. (zaicongwang@gmail.com) or H.B. (hbecker@zedat.fu-berlin.de).

METHODS

Concentrations of S, Se, Te and the HSEs in bulk rocks of peridotites and some carbonaceous chondrites (Supplementary Information) were determined by isotope dilution and inductively coupled plasma mass spectrometry using HPA-S digestion. Details of the techniques have been described elsewhere^{16,23}. Additional data on duplicates of peridotites, reference materials and some carbonaceous chondrites are presented in Supplementary Tables 1–3. About 2.5 g of peridotite powder or 50–70 mg of meteorite sample powder were digested by HPA-S in 5 ml of 14 M HNO₃ and 2.5 ml of 9 M HCl at 100 bar and 320 °C for 16 h after addition of a mixed ⁷⁷Se–¹²⁵Te and a ³⁴S spike solution for all samples, and a mixed ¹⁹¹Ir–⁹⁹Ru–¹⁹⁴Pt–¹⁰⁵Pd spike for some samples to obtain duplicate data for comparison with literature values. After digestion, one third of the digestion solution was used for HSE chemical separation^{16,23} (Supplementary Table 1) and about one-quarter of the remaining solution was used for S–Se–Te separation.

Chemical separation. A two-step anion- and cation-exchange chromatography method was used for the separation of S–Se from Te and matrix elements. The digestion solution was evaporated, repeatedly treated with 9 M HCl and dried to reduce Se and Te to the quadrivalence state. After dissolution of the residue in 4 ml of 6 M HCl, the solution was loaded on 3 ml of pre-cleaned Eichrom 1-X8 (100–200 mesh) anion resin^{23,31}. Sulphur and Se and part of the matrix were collected in 7 ml of 6 M HCl. The resin was further washed with 5 ml of 5 M hydrofluoric acid and 4 ml of 9 M HCl to elute remaining matrix. Finally, Te was eluted in 7 ml of 1 M HNO₃, ready for mass spectrometric analysis. The S–Se fraction was dried at 85 °C and the residue dissolved in 4 ml of 0.1 M HNO₃. The S–Se fraction was purified on 3 ml of Eichrom 50W-X8 (100–200 mesh) cation resin and collected in 7 ml of 0.1 M HNO₃. We added 1 ml of 7 M HNO₃ to facilitate the hydride generation reaction. Sulphur and Se show high recovery (>80%), but Te yields were variable, ranging from 10% to >95% depending on the composition of the matrix: they were typically around 10–40% for peridotites, >80% for chondrites and >95% for procedural blanks. The low Te yield for peridotites is probably due

to incomplete reduction of Te. Yields of Te can be improved by heating the S–Se fraction at 85 °C for two hours and reloading on the anion resin after washing the resin in 5 M hydrofluoric acid and 9 M HCl. Precise and accurate Te data can be obtained using the isotope dilution method and hydride generation²³.

Inductively coupled plasma mass spectrometry. Sulphur isotopic measurements were performed on the S–Se fraction using a Scott-type glass spray chamber on the Element XR ICP-MS at Freie Universität Berlin in medium-mass-resolution mode ($m/\Delta m = 4,000$). Precisions of ³²S/³⁴S were greater than $\pm 0.1\%$ (twice the relative standard error (2 r.s.e.)) using $n = 800$ scans and a sample time of 10 ms. Selenium and Te were measured in low-mass-resolution mode on the Element XR, using a double-pass Scott-type glass spray chamber, combined with a hydride-generation sample-introduction system, by reacting the sample solution with 1% m/m NaBH₄ in 0.05 M NaOH (ref. 23). Owing to variations in the gas flow during hydride generation, the Se and Te intensities were not as stable as for the HSEs and S, with a typical uncertainty of 3–8% relative standard deviations (r.s.d.) for samples. Standard errors of the mean of ⁷⁷Se/⁸²Se and ¹²⁵Te/¹²⁶Te, which were used for concentration calculation, were less than 0.1% (2 r.s.e.) using sample times of 5 ms and $n = 6,000$ and 9,000 scans, respectively. ⁷⁷Se/⁸²Se and ¹²⁵Te/¹²⁶Te ratios of Se and Te standard solutions changed little (typically 0.1% to <1%) during the measurement period, leading to variations in Se and Te concentrations of <1–2%.

Procedural blanks. The S blank was typically $3 \pm 1.5 \mu\text{g}$ (1 r.s.d., $n = 11$; mainly from the cation resin), the Se blank was typically $1.5 \pm 1 \text{ ng}$ (1 r.s.d., $n = 10$) and the Te blanks were typically $20 \pm 15 \text{ pg}$ (1 r.s.d., $n = 10$) during the analysis period. Procedural-blank corrections for the chalcogens were always applied and were in most cases insignificant (<1%). For depleted harzburgites, procedural blanks may comprise up to 5–25% for S, Se and Te when using 2.5 g of sample powder.

31. Fehr, M. A., Rehkämper, M. & Halliday, A. N. Application of MC-ICPMS to the precise determination of tellurium isotope compositions in chondrites, iron meteorites and sulfides. *Int. J. Mass Spectrom.* **232**, 83–94 (2004).

Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise

Trevor F. Keenan¹, David Y. Hollinger², Gil Bohrer³, Danilo Dragoni⁴, J. William Munger⁵, Hans Peter Schmid⁶ & Andrew D. Richardson¹

Terrestrial plants remove CO₂ from the atmosphere through photosynthesis, a process that is accompanied by the loss of water vapour from leaves¹. The ratio of water loss to carbon gain, or water-use efficiency, is a key characteristic of ecosystem function that is central to the global cycles of water, energy and carbon². Here we analyse direct, long-term measurements of whole-ecosystem carbon and water exchange³. We find a substantial increase in water-use efficiency in temperate and boreal forests of the Northern Hemisphere over the past two decades. We systematically assess various competing hypotheses to explain this trend, and find that the observed increase is most consistent with a strong CO₂ fertilization effect. The results suggest a partial closure of stomata¹—small pores on the leaf surface that regulate gas exchange—to maintain a near-constant concentration of CO₂ inside the leaf even under continually increasing atmospheric CO₂ levels. The observed increase in forest water-use efficiency is larger than that predicted by existing theory and 13 terrestrial biosphere models. The increase is associated with trends of increasing ecosystem-level photosynthesis and net carbon uptake, and decreasing evapotranspiration. Our findings suggest a shift in the carbon- and water-based economics of terrestrial vegetation, which may require a reassessment of the role of stomatal control in regulating interactions between forests and climate change, and a re-evaluation of coupled vegetation–climate models.

Plants assimilate atmospheric CO₂ through photosynthesis. The complex organic molecules that result ultimately support most of life on Earth. Acquisition of CO₂ is accompanied by the loss of water vapour through stomata, the small pores on leaf surfaces that regulate the diffusion of these two gases between the leaf and the atmosphere^{1,4,5}. The rate of carbon uptake per unit of water lost, referred to as water-use efficiency, integrates a suite of biotic and abiotic factors, and, importantly, quantifies how much water an ecosystem uses relative to carbon gained.

The photosynthetic uptake of CO₂, A , and the loss of water vapour from a leaf to the atmosphere (transpiration, E) can both be described as a conductance g_s multiplied by a concentration difference between the atmosphere and the internal leaf mesophyll:

$$A = g_s(c_a - c_i) \quad (1)$$

$$E = 1.6g_s(v_i - v_a) \quad (2)$$

where g_s is the stomatal conductance to CO₂, c_a and c_i represent the ambient and leaf intercellular mole fractions of CO₂ (respectively), v_a and v_i similarly represent the ambient and leaf intercellular water vapour mole fractions, and 1.6 accounts for the greater diffusivity of water vapour relative to CO₂.

At the leaf level, instantaneous water-use efficiency W quantifies the rate of carbon uptake per unit of water lost ($W = A/E$). Similarly, at the ecosystem scale, bulk canopy water-use efficiency is defined as:

$$W_e = \frac{A_e}{E_e} \quad (3)$$

To account for the effects of atmospheric evaporative demand, D , an ‘inherent’ ecosystem-scale water-use efficiency² can be calculated as $W_{ei} = W_e D$ (see Methods). Combining W_{ei} with equations (1) and (2) shows that the sensitivity of W_{ei} to c_a depends on how c_i varies with c_a , and:

$$\Delta W_{ei} = \Delta(c_a - C_i)/1.6 \quad (4)$$

where C_i represents the integrated canopy leaf intercellular mole fraction of CO₂ and Δ is the difference operator.

The biochemistry of photosynthesis indicates that increasing atmospheric CO₂ should lead to an increase in both photosynthetic uptake and water-use efficiency⁴—that is, the ‘CO₂ fertilization effect’. This could stimulate carbon uptake and storage by the terrestrial biosphere, and may account for largely unexplained increases in forest growth rates^{6–8} and terrestrial carbon uptake⁹. CO₂-enrichment experiments¹⁰, proxy reconstructions¹¹, and models provide support for an increase in plant resource-use efficiency as CO₂ concentrations increase, but are often inconsistent and inconclusive^{10,12–14}. The magnitude of any response to elevated CO₂ in natural forest ecosystems is largely unknown^{15,16}, leading to debate on the effect of nutrient limitations¹⁷. This is not surprising, given the difficulties in translating results from isotopic proxies and small-scale manipulative experiments to intact natural ecosystems^{10,12}.

Here we document and analyse recent trends in the inherent water-use efficiency W_{ei} of forest canopies using direct and continuous long-term measurements of CO₂ and water vapour fluxes. We address a variety of factors as possible drivers of the observed trends, including the potential role of increasing levels of atmospheric CO₂. We focus on seven sites in the midwestern and northeastern United States. The forests represent tree species compositions typical of the region, and are not actively managed. This regional analysis is then compared with data from 14 additional temperate and boreal forest sites (FLUXNET, www.fluxdata.org) more widely distributed across the Northern Hemisphere. We use the half-hourly eddy-covariance flux measurements, and concurrently recorded meteorological data, to calculate W_{ei} as the ratio of canopy carbon assimilation to ecosystem evapotranspiration, adjusted for atmospheric evaporative demand (see Methods).

The regional analysis shows that W_{ei} has increased at each forest site over the measurement period (Fig. 1, Supplementary Table 2). Of the seven regional focus sites, the change was significant ($P < 0.05$) at three sites, whereas at all others the slope was positive but with lower P values. The mean trend (± 1 s.e.) across all seven sites is 1.07 ± 0.3

¹Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, 02138, USA. ²USDA Forest Service, Northern Research Station, Durham, New Hampshire 03824, USA. ³Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, Columbus, Ohio 43210, USA. ⁴Department of Geography, Indiana University, Bloomington, Indiana 47405, USA. ⁵School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. ⁶Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology, IMK-IFU, 82467 Garmisch-Partenkirchen, Germany.

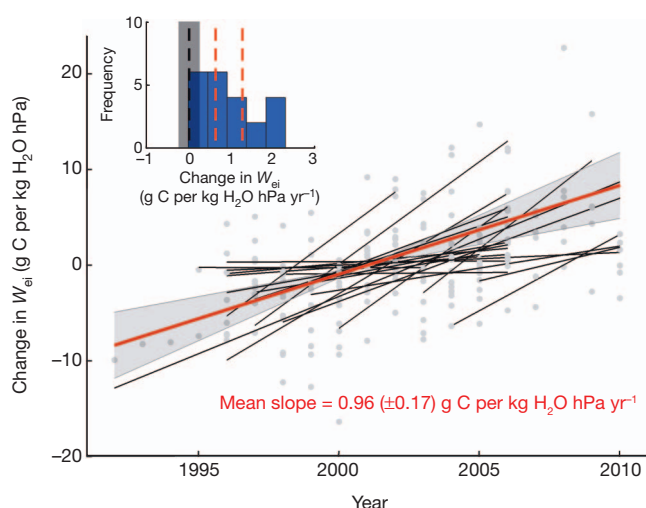


Figure 1 | Long-term change in forest water-use efficiency. The annual change in water use efficiency, ΔW_{ei} (g C per kg H₂O hPa; here, evaporative demand is expressed as vapour-pressure deficit), calculated using daytime fluxes from summer months at all sites. ΔW_{ei} is normalized to mean W_{ei} at each site. The red line represents the mean trend over all sites, extrapolated over the entire measurement period. The grey area highlights trends within one standard deviation about the mean trend. Individual site observations and trends are given as grey dots and black lines respectively. Individual site names, details, statistics and plots are given in the Supplementary Information. The inset shows the distribution of the rate of change in W_{ei} over all sites. Dashed red lines represent bootstrapped 95% confidence intervals for the mean slope. The grey shaded area represents the mean and standard deviation of the slope of increases in W_{ei} simulated by 13 terrestrial biosphere models at four of the US regional sites. See Supplementary Information for more information on the model runs and individual site graphs.

grams of carbon per kilogram of water times a hectopascal per year (g C per kg H₂O hPa yr⁻¹) ($P = 0.011$, Student's t -test), with a joint probability of $P = 0.0016$ (Fisher combined probability). This trend is replicated in the broader FLUXNET data set as well. Eight of these 14 sites show an increase in W_{ei} over the period of record with $P < 0.15$ (Fig. 1), and three of the six remaining sites show increases in W_{ei} with $P < 0.3$ (Supplementary Table 2); for the FLUXNET sites, the joint probability of a significant trend is $P = 1.1 \times 10^{-4}$ (Fisher combined probability). Thus, although the trend is not statistically significant at all individual sites, it is highly significant when multiple sites are considered. For the larger FLUXNET data set, the mean trend is 0.90 ± 0.2 g C per kg H₂O hPa yr⁻¹ ($P < 0.001$, t -test), which is not significantly different from the regional analysis sites. Over all sites, deciduous forest sites show a significantly higher ($P = 0.06$, t -test) rate of response (1.4 ± 0.2 , $n = 8$) than do evergreen forest sites (0.64 ± 0.3 , $n = 10$, mixed forests excluded).

In addition to rising CO₂, a number of other factors, both environmental and biotic, could cause the observed increase in forest water-use efficiency. These include: climate change; nitrogen deposition and accumulation; changes in leaf area, canopy height, surface roughness and the coupling of the canopy to the atmosphere; and long-term instrument drift. We tested each of these competing hypotheses using data from the regional focus sites. To examine whether the increase in W_{ei} can be linked with recent changes in environmental factors affecting plant water use, we stratified the observations for different measurement conditions and tested for trends in climatic variables across all sites (see Supplementary Information sections 6 and 8). This analysis shows that the general trend and magnitude of the increase in W_{ei} is independent of recent changes in environmental factors. The lack of a consistent trend in climatic forcing across all sites (with the exception of vapour pressure deficit, D) further suggests that changes in climatic drivers are not responsible for the consistent increase in W_{ei} .

Similarly, we analysed long-term measurements of leaf area and leaf nitrogen content (where available), together with estimates of surface roughness, and found no significant changes over time (see Supplementary Information sections 9–11). The lack of trends in surface roughness and leaf area, and the lack of a decrease in measured above-canopy wind speeds, suggests that the coupling of the canopy to the atmosphere has remained relatively constant during the measurement period. To minimize instrument biases, strong quality controls are in place at each site¹⁸. These include *in situ* comparisons with a roving standard instrument, and regular changes of inlet tubes and filters. The variation in magnitude of trends in W_{ei} across sites suggests that local factors (for example, droughts) could result in site-specific trends that are superimposed on the more general response of forests worldwide. Our detailed analysis (see Supplementary Information) suggests that of all the potential drivers of the observed changes in W_{ei} , the only driver that is changing sufficiently and consistently through time at all sites is atmospheric CO₂.

To examine whether the rate of increase in water-use efficiency is consistent with a fertilization effect of elevated atmospheric CO₂, c_a , we combined the flux measurements, observed c_a , and equations (1) to (4), to estimate C_i , the canopy-level equivalent to leaf c_i (see equation (9) in the Methods). This analysis suggests that, if rising CO₂ is the driving factor, then to reproduce the magnitude of the observed trends in W_{ei} , plants would have to partially close their stomata (thereby reducing g_s and hence E) to hold C_i at a relatively constant level (Fig. 2) despite increasing c_a concentrations. By comparison, current theoretical stomatal control models^{5,19,20} project a relatively limited enhancement of W_{ei} under elevated c_a . Indeed, we show that 13 state-of-the-art terrestrial biosphere models do not predict long-term trends in W_{ei} (0.02 ± 0.25 g C per kg H₂O hPa yr⁻¹) for forested sites in North America (Fig. 1), in contrast to the observational data presented here. Although the coupling of c_i to c_a is commonly reported to be proportional¹⁴, recent studies report that a near-constant level of c_i under elevated CO₂ is indeed plausible¹⁶.

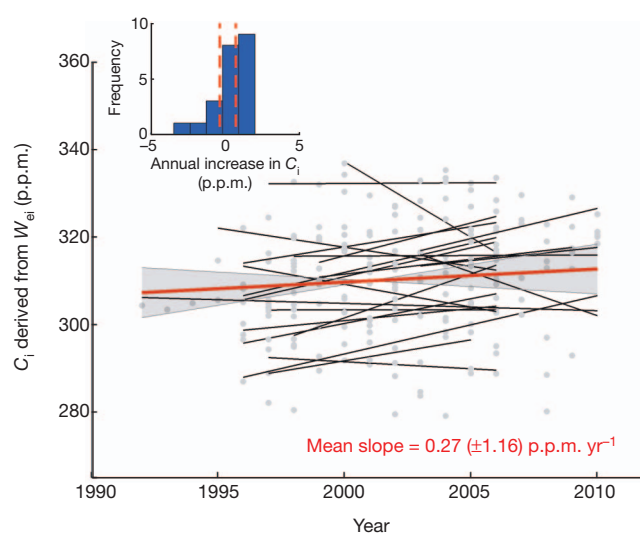


Figure 2 | Canopy-scale leaf intercellular CO₂ concentrations. The derived canopy-scale intercellular CO₂ concentration C_i at all sites. The red line represents the mean trend over all sites, extrapolated over the entire measurement period. The grey area highlights trends within one standard deviation about the mean trend. Individual site observations and trends are given as grey dots and black lines respectively. Individual site names, details, and plots are given in the Supplementary Information. The inset shows the distribution of the rate of change in C_i over all sites. Dashed red lines represent bootstrapped 95% confidence intervals for the mean slope. The distribution of slopes around zero indicates that the observed changes in W_{ei} are consistent with a constant C_i . Site details are given in Supplementary Table 1.

The observed trends in W_{ei} have important consequences for water and carbon cycling in forest ecosystems. The direct tradeoff between water loss and carbon uptake through the stomata—equations (1) and (2)—means that, as water-use efficiency increases, either evapotranspiration (E_e) decreases or gross photosynthetic carbon uptake increases, or both occur simultaneously. Mid-summer forest E_e declined at all but one of the regional US sites. Decreases in E_e are consistent with recent declines in evapotranspiration, and increases in streamflow, at watersheds in the northeastern US²¹. In the global FLUXNET data set, ten of the 14 sites also show a trend of reduced E_e (Supplementary Table 2). Three of the remaining 4 FLUXNET sites that did not demonstrate a decrease in E_e had increased rates of summer photosynthesis, consistent with the overall increase in W_{ei} (Supplementary Table 2) and the water–carbon tradeoff. The mean annual trend in E_e across all sites was $-3.8 \pm 2 \text{ g H}_2\text{O m}^{-2} \text{ h}^{-1} \text{ yr}^{-1}$ ($P = 0.07$, t -test).

Additionally, mid-summer daytime forest net carbon uptake increased during the measurement period for six of the seven regional US sites ($13.1 \pm 5.6 \text{ mg C m}^{-2} \text{ h}^{-1} \text{ yr}^{-1}$, $P = 0.03$, t -test). At those regional sites with sufficient data to construct annual totals, we observed strong trends in total annual net carbon uptake (Fig. 3). In the FLUXNET analysis, only six of the 14 sites showed trends of increased net carbon uptake over the measurement period, with probabilities ranging between $P = 0.01$ and 0.59 (Supplementary Table 2). The mean trend across all 21 sites was $6.8 \pm 3.6 \text{ mg C m}^{-2} \text{ h}^{-1} \text{ yr}^{-1}$ ($P = 0.06$, t -test).

To further examine trends in carbon uptake, we carried out a model-data fusion analysis using a parsimonious process-based model (see Supplementary Information section 7). This analysis indicates that only a small fraction of trends in carbon uptake can be explained by changes in climate forcings (temperature, precipitation, humidity, solar radiation) at any site. Previous work attributes a proportion of the net annual increase to a lengthening of the growing season²². Our model-data fusion approach, however, shows a large increase in net uptake during the summer months, independent of season length.

The observed increase in water-use efficiency documented here has a range of important implications for ecosystem function, services and feedbacks to the climate system. These include enhanced timber yields¹⁰, improved water availability (which could partially offset the effects of future droughts), and changes in competitive interactions²³. E_e directly

affects the surface energy balance. Reduced E_e due to higher water-use efficiency²⁴ could therefore lead to higher air temperatures²⁵, decreased humidity, and decreased recycling of continental precipitation²⁶. This would give rise to increased continental freshwater runoff²⁶, along with drought in parts of the world that rely on water transpired in other regions²⁷. Increases in W_{ei} may account for reports of global increases in photosynthesis²⁸, forest growth rates^{6–8}, and carbon uptake⁹. Our analysis suggests that rising atmospheric CO_2 is having a direct and unexpectedly strong influence on ecosystem processes and biosphere–atmosphere interactions in temperate and boreal forests. Understanding how increasing CO_2 induces shifts in terrestrial carbon uptake and water loss and long-term changes in water-use efficiency is of critical importance for improving our ability to project the future evolution of the Earth system.

METHODS SUMMARY

Our analysis is based on half-hourly data from 21 flux towers distributed across northern temperate and boreal forest ecosystems. Measurements used include canopy-scale water vapour flux, CO_2 flux, meteorological variables, and estimates of gross primary photosynthesis derived from the CO_2 flux measurements using the eddy covariance technique³. We used two distinct data sets, representing both a regional and global focus. The first data set comprised seven unmanaged forested sites in the midwestern and northeastern United States (Supplementary Table 1). The second data set comprised 14 additional forest sites, widely distributed throughout the Northern Hemisphere and subject to various management regimes. The data records ranged in length from 7 to 18 years. We estimated canopy water-use efficiency (W_{ei}) as the ratio of the hourly sums of summer-month daytime gross ecosystem photosynthesis to ecosystem transpiration, adjusted for atmospheric evaporative demand². Trends were estimated using the Mann–Kendall Tau non-parametric trend test with Sen's method, and significance levels on the basis of Mann–Kendall tests.

Full Methods and any associated references are available in the online version of the paper.

Received 13 September 2012; accepted 13 May 2013.

Published online 10 July 2013.

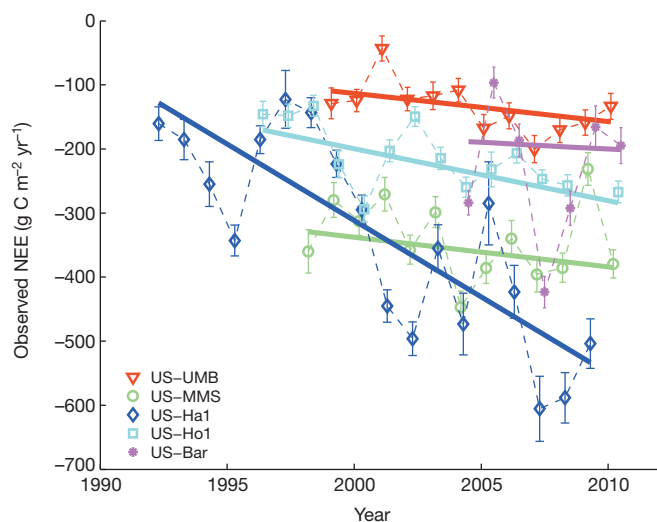


Figure 3 | Long-term increase in net ecosystem carbon uptake. The annual net ecosystem carbon uptake (NEE) at five natural forest sites in the northeastern USA. Solid lines represent sen-slope estimates (Methods). Error bars represent the 95% confidence interval, derived by adding uncertainty due to random measurement error and gap-filling uncertainty, using 1,000 bootstrap samples. For two sites included in the regional database there were too many gaps to determine reliable annual sums. Site details are given in Supplementary Table 1.

- Berry, J. A., Beerling, D. J. & Franks, P. J. Stomata: key players in the earth system, past and present. *Curr. Opin. Plant Biol.* **13**, 232–239 (2010).
- Beer, C. *et al.* Temporal and among-site variability of inherent water use efficiency at the ecosystem level. *Glob. Biogeochem. Cycles* **23**, 1–13 (2009).
- Baldocchi, D. Breathing of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Aust. J. Bot.* **56**, 1–26 (2008).
- Field, C. B., Jackson, R. B. & Mooney, H. A. Stomatal responses to increased CO_2 : implications from the plant to the global scale. *Plant Cell Environ.* **18**, 1214–1225 (1995).
- Katul, G. G., Palmroth, S. & Oren, R. Leaf stomatal responses to vapour pressure deficit under current and CO_2 -enriched atmosphere explained by the economics of gas exchange. *Plant Cell Environ.* **32**, 968–979 (2009).
- Lewis, S. L. *et al.* Increasing carbon storage in intact African tropical forests. *Nature* **457**, 1003–1006 (2009).
- Salzer, M. W., Hughes, M. K., Bunn, A. G. & Kipfmüller, K. F. Recent unprecedented tree-ring growth in bristlecone pine at the highest elevations and possible causes. *Proc. Natl Acad. Sci. USA* **106**, 20348–20353 (2009).
- McMahon, S. M., Parker, G. G. & Miller, D. R. Evidence for a recent increase in forest growth. *Proc. Natl Acad. Sci. USA* **107**, 3611–3615 (2010).
- Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P. & White, J. W. C. Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature* **488**, 70–72 (2012).
- Norby, R. J. & Zak, D. R. Ecological lessons from free-air CO_2 enrichment (FACE) experiments. *Annu. Rev. Ecol. Evol. Syst.* **42**, 181–203 (2011).
- Peñuelas, J., Canadell, J. G. & Ogaya, R. Increased water-use efficiency during the 20th century did not translate into enhanced tree growth. *Glob. Ecol. Biogeogr.* **20**, 597–608 (2011).
- Seibt, U., Rajabi, A., Griffiths, H. & Berry, J. A. Carbon isotopes and water use efficiency: sense and sensitivity. *Oecologia* **155**, 441–454 (2008).
- Schubert, B. A. & Jahren, A. H. The effect of atmospheric CO_2 concentration on carbon isotope fractionation in C3 land plants. *Geochim. Cosmochim. Acta* **96**, 29–43 (2012).
- Ainsworth, E. A. & Rogers, A. The response of photosynthesis and stomatal conductance to rising $[\text{CO}_2]$: mechanisms and environmental interactions. *Plant Cell Environ.* **30**, 258–270 (2007).
- Saurer, M., Siegwolf, R. T. W. & Schweingruber, F. H. Carbon isotope discrimination indicates improving water-use efficiency of trees in northern Eurasia over the last 100 years. *Glob. Change Biol.* **10**, 2109–2120 (2004).

16. Battipaglia, G. *et al.* Elevated CO₂ increases tree-level intrinsic water use efficiency: insights from carbon and oxygen isotope analyses in tree rings across three forest FACE sites. *New Phytol.* **197**, 544–554 (2013).
17. Drake, J. E. *et al.* Increases in the flux of carbon belowground stimulate nitrogen uptake and sustain the long-term enhancement of forest productivity under elevated CO₂. *Ecol. Lett.* **14**, 349–357 (2011).
18. Schmidt, A., Hanson, C., Chan, W. S. & Law, B. E. Empirical assessment of uncertainties of meteorological parameters and turbulent fluxes in the AmeriFlux network. *J. Geophys. Res.* **117**, G04014 (2012).
19. Medlyn, B. E. *et al.* Reconciling the optimal and empirical approaches to modelling stomatal conductance. *Glob. Change Biol.* **17**, 2134–2144 (2011).
20. Buckley, T. N. The role of stomatal acclimation in modelling tree adaptation to high CO₂. *J. Exp. Bot.* **59**, 1951–1961 (2008).
21. Jones, J. A. *et al.* Ecosystem processes and human influences regulate streamflow response to climate change at long-term ecological research sites. *Bioscience* **62**, 390–404 (2012).
22. Dragoni, D. *et al.* Evidence of increased net ecosystem productivity associated with a longer vegetated season in a deciduous forest in south-central Indiana, USA. *Glob. Change Biol.* **17**, 886–897 (2011).
23. Niinemets, U., Flexas, J. & Peñuelas, J. Evergreens favored by higher responsiveness to increased CO₂. *Trends Ecol. Evol.* **26**, 136–142 (2011).
24. Bernacchi, C. J., Kimball, B. A., Quarles, D. R., Long, S. P. & Ort, D. R. Decreases in stomatal conductance of soybean under open-air elevation of [CO₂] are closely coupled with decreases in ecosystem evapotranspiration. *Plant Physiol.* **143**, 134–144 (2007).
25. Lee, X. *et al.* Observed increase in local cooling effect of deforestation at higher latitudes. *Nature* **479**, 384–387 (2011).
26. Betts, R. A. *et al.* Projected increase in continental runoff due to plant responses to increasing carbon dioxide. *Nature* **448**, 1037–1041 (2007).
27. Keys, P. W. *et al.* Analyzing precipitation sheds to understand the vulnerability of rainfall dependent regions. *Biogeosciences* **9**, 733–746 (2012).
28. Nemani, R. R. *et al.* Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* **300**, 1560–1563 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements This research was supported by the NOAA Climate Program Office, Global Carbon Cycle Program (award NA11OAR4310054) and the Office of

Science (Biological and Environmental Research), US Department of Energy. G.B. acknowledges a grant from the National Science Foundation (grant number DEB-0911461). This work used eddy covariance data acquired by the FLUXNET community and in particular by the AmeriFlux, CarboEuropeIP and Fluxnet-Canada networks. AmeriFlux was supported by the US Department of Energy, Biological and Environmental Research, Terrestrial Carbon Program (grant numbers DE-FG02-04ER63917 and DE-FG02-04ER63911, DE-SC0006708) and Fluxnet-Canada was supported by CFCAS, NSERC, BIOCAP, Environment Canada and NRCan. We acknowledge financial support of the eddy covariance data harmonization provided by CarboEuropeIP, FAO-GTOS-TCO, iLEAPS, the Max-Planck Institute for Biogeochemistry, National Science Foundation, University of Tuscia, Université Laval and Environment Canada and US Department of Energy and of the database development and technical support from Berkeley Water Center, Lawrence Berkeley National Laboratory, Microsoft Research eScience, Oak Ridge National Laboratory, University of California-Berkeley, University of Virginia. We thank all those involved in the NACP Site Synthesis, in particular the modelling teams who provided model output. Research at the Bartlett Experimental Forest tower is supported by the National Science Foundation (grant DEB-1114804), and the USDA Forest Service's Northern Research Station. Research at Howland Forest is supported by the Office of Science (BER), US Department of Energy. Carbon flux and biometric measurements at Harvard Forest have been supported by the Office of Science (BER), US Department of Energy and the National Science Foundation Long-Term Ecological Research programmes. We thank S. Ollinger and S. Frey for maintaining the long-term leaf nitrogen measurements at Harvard Forest, and B. Yang for providing gap-filled meteorological data for the regional focus sites.

Author Contributions T.F.K. and A.D.R. designed the study and are responsible for the integrity of the manuscript. A.D.R. planned the regional analysis, with input from D.Y.H., J.W.M., G.B., H.P.S. and D.D. A.D.R., D.Y.H., J.W.M., G.B., H.P.S. and D.D. contributed data. T.F.K. compiled the data sets, detailed and performed the analysis. A.D.R. and D.Y.H. contributed ideas to the analysis. T.F.K. led the writing, with input from A.D.R. and D.Y.H. All authors discussed and commented on the results and the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.F.K. (tkeenanoeb@oeb.harvard.edu) or A.D.R. (arichardson@oeb.harvard.edu).

METHODS

We used high-frequency (30 min) measurements of water vapour flux, CO₂ flux, meteorological variables and estimates of gross primary photosynthesis derived from the CO₂ flux measurements. Flux measurements were made at the ecosystem level based on the eddy covariance technique, a method that allows the direct measurement of fluxes of CO₂ and water vapour between Earth's surface and the atmosphere³. Two distinct data sets were used, representing both a regional and global focus. The first data set comprises seven unmanaged forested sites in the midwestern and northeastern United States (Supplementary Table 1), suitable for characterizing large-scale regional tendencies. Flux measurements at these sites are all ongoing. The data records we analyse here range in length from seven to 18 years. The second data set (La Thuile FLUXNET 'fair use' database; www.fluxdata.org) consists of a freely available global collection of eddy covariance CO₂ and water vapour flux observations for a variety of ecosystem types. From this database we selected forest sites with closed canopies that had seven years or more of measurements (see criteria below). A description of the seven regional focus sites and the 14 La Thuile sites used in this study can be found in Supplementary Table 1. At all sites, CO₂ and H₂O fluxes were measured using a closed path eddy-covariance system with an LI-6262 gas analyser (changed to LI-7000 at US-MMS in 2004, US-Ha2 in 2004, FI-Sod in 2003, and US-UMB at 2006). Regular tube and filter replacement were performed to ensure high-quality water flux data, along with detailed Ameriflux quality control campaigns¹⁸.

Site selection from the FLUXNET La Thuile data set. We initially considered all freely available data from the 'fair use' La Thuile FLUXNET data set. Of the 153 sites available, we selected the 31 sites that had seven or more years of flux measurements. We focused our analysis on sites that represented temperate and boreal forests. This meant the exclusion of two wetlands, one cropland, one grassland and seven Mediterranean forests. Mediterranean forests are dominated by interannual variability in soil water availability, which complicates the derivation of changes in water-use efficiency driven by atmospheric CO₂. Because water-use efficiency is affected by changes in forest canopy height, we also excluded data from two temperate forests that were recently planted (<25 years old). One eligible forested site (DE-Hai) was excluded owing to errors in water vapour flux estimates that scaled with humidity (La Thuile database curators, personal communication.). This left a total of 14 temperate and boreal forests in the 'fair use' La Thuile data set that met our selection criteria (excluding those that are already in the regional US data set). Note that, in contrast to the seven regional US sites, many of these forests are subject to some level of management.

Flux data processing. Gross ecosystem photosynthesis (GEP) represents the carboxylation rate minus photorespiration in this study. At night, NEE consists of all respiratory processes except photorespiration. Accounting for the temperature sensitivity of this respiration component results in the equivalent respiration during the day which is further subtracted from daytime NEE to derive GEP²⁹. There are a variety of approaches to derive GEP³⁰. Previous comparisons have shown good agreement between different approaches but recommend the consistent use of a particular approach across sites³⁰. Common data processing was performed for all sites. Carbon fluxes were corrected for the storage component and spikes³¹. Data with low turbulent mixing were filtered out using a threshold for friction velocity²⁹. GEP for all sites was estimated using the standard FLUXNET on-line flux-partitioning tool (www.bgc-jena.mpg.de/~MDIwork/eddyproc/).

Derivation of canopy-integrated water use efficiency and intercellular CO₂ concentrations. For whole plants or ecosystems, water-use efficiency (W_e) can be calculated in a similar fashion as for leaves, as:

$$W_e = \frac{\text{GEP}}{E_c(1 - \phi_w)} \quad (5)$$

where E_c is ecosystem evapotranspiration, and ϕ_w represents the fraction of non-transpiratory water loss.

Transpiration, and thus W_e , are a function of evaporative demand D . 'Inherent' water-use efficiency (W_{ei}) is therefore used when comparing water-use efficiency between species or meteorological conditions^{2,32}. At the ecosystem level, W_{ei} can be approximated using eddy-covariance flux measurements² (given certain assumptions outlined below), as the ratio between GEP and canopy conductance G_s :

$$W_{ei} = \text{GEP}/G_s \quad (6)$$

To do so, both GEP and G_s must first be derived from the measured carbon and water fluxes. GEP was estimated for each site with commonly used flux partitioning algorithms (see above). G_s can be estimated from latent heat measurements given the following assumptions: (1) vapour pressure difference between the leaf and the atmosphere can be approximated by measured atmospheric evaporative demand (D), assuming equal temperatures of leaves and atmosphere, (2) aerodynamic resistance between the canopy and the reference-height for the flux can be neglected, (3) under dry conditions, with no recent precipitation events, measured water vapour fluxes are equivalent to transpiration, giving $\phi_w = 0$ (that is, evaporation contributes minimally). These assumptions allow the definition of the equality $E_c(1 - \phi_w) = G_s D$.

An ecosystem-level representation of water-use efficiency can then be formulated as²:

$$W_{ei} = \text{GEP} \times D/E_c \quad (7)$$

and

$$W_{ei} = (c_a - C_i)/1.6 \quad (8)$$

where C_i is the canopy integrated concentration of intercellular CO₂, and 1.6 is a physical constant that accounts for the greater diffusivity of water vapour relative to CO₂. We thus calculated W_{ei} directly from flux measurements of water vapour (latent heat flux), meteorological measurements of D , and CO₂ flux-based estimates of GEP, as the ratio of the sum of the filtered fluxes each year. We used hourly or half-hourly measurements for the summer months of June, July and August at each site. Only measured values were used (that is, no gap filled data). Night-time data were excluded, because leaves are photosynthetically inactive at night. During and after rain events, evaporation from wet leaves can contribute largely to measured water fluxes. To focus on stomatal controlled water fluxes, days with rain events, and the day after a rain event, were excluded from the analysis, as interception storage is largely depleted within two days following rain events³³. Soil evaporation below closed-forest canopies is considered negligible. The W_{ei} values reported here therefore represented mean values for summer daytime hours. See ref. 2 for further details of the derivation and justification of W_{ei} , and an analysis of its spatial variability.

Given the atmospheric CO₂ concentrations (www.esrl.noaa.gov/gmd/), equation (8) can be inverted to estimate C_i . We used the flux-derived W_{ei} , and atmospheric CO₂, to estimate C_i as:

$$C_i = c_a - 1.6W_{ei} \quad (9)$$

It should be noted that the magnitude of W_{ei} at each site could be biased owing to the assumptions applied, and owing to systematic biases in the derived GEP, because it is not a directly measurable flux. Moreover, the available estimates of c_a are regional values and may slightly deviate from local canopy-scale values, thus contributing to potential site-biases. Such biases, however, are not expected to change over long timescales, given that the flux time series at each site is partitioned to GEP using a common algorithm (www.bgc-jena.mpg.de/~MDIwork/eddyproc/). Any biases in W_{ei} or c_a will, however, translate to biases in the magnitude of the derived C_i . This could lead to between-site differences in the mean C_i that are not directly interpretable in an ecological context.

29. Reichstein, M. *et al.* On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Glob. Change Biol.* **11**, 1424–1439 (2005).
30. Desai, A. R. *et al.* Cross-site evaluation of eddy covariance GPP and RE decomposition techniques. *Agric. For. Meteorol.* **148**, 821–838 (2008).
31. Papale, D. *et al.* Towards a standardized processing of net ecosystem exchange measured with eddy covariance technique: algorithms and uncertainty estimation. *Biogeosciences* **3**, 571–583 (2006).
32. Khairi, M. M. A. & Hall, A. E. Comparative studies of net photosynthesis and transpiration of some citrus species and relatives. *Physiol. Plant.* **36**, 35–39 (1976).
33. Grelle, A., Lundberg, A., Lindroth, A., Morén, A.-S. & Cienciala, E. Evaporation components of a boreal forest: variations during the growing season. *J. Hydrol.* **197**, 70–87 (1997).

Elucidation of the Fe(IV)=O intermediate in the catalytic cycle of the halogenase SyrB2

Shaun D. Wong^{1*}, Martin Srncic^{1*}, Megan L. Matthews^{2†}, Lei V. Liu¹, Yeonju Kwak¹, Kiyoun Park¹, Caleb B. Bell III¹, E. Ercan Alp³, Jiyong Zhao³, Yoshitaka Yoda⁴, Shinji Kitao⁵, Makoto Seto⁵, Carsten Krebs^{2,6}, J. Martin Bollinger Jr^{2,6} & Edward I. Solomon^{1,7}

Mononuclear non-haem iron (NHF) enzymes catalyse a broad range of oxidative reactions, including halogenation, hydroxylation, ring closure, desaturation and aromatic ring cleavage reactions. They are involved in a number of biological processes, including phenylalanine metabolism, the production of neurotransmitters, the hypoxic response and the biosynthesis of secondary metabolites^{1–3}. The reactive intermediate in the catalytic cycles of these enzymes is a high-spin $S = 2$ Fe(IV)=O species, which has been trapped for a number of NHF enzymes^{4–8}, including the halogenase SyrB2 (syringomycin biosynthesis enzyme 2). Computational studies aimed at understanding the reactivity of this Fe(IV)=O intermediate^{9–13} are limited in applicability owing to the paucity of experimental knowledge about its geometric and electronic structure. Synchrotron-based nuclear resonance vibrational spectroscopy (NRVS) is a sensitive and effective method that defines the dependence of the vibrational modes involving Fe on the nature of the Fe(IV)=O active site^{14–16}. Here we present NRVS structural characterization of the reactive Fe(IV)=O intermediate of a NHF enzyme, namely the halogenase SyrB2 from the bacterium *Pseudomonas syringae* pv. *syringae*. This intermediate reacts via an initial hydrogen-atom abstraction step, performing subsequent halogenation of the native substrate or hydroxylation of non-native substrates¹⁷. A correlation of the experimental NRVS data to electronic structure calculations indicates that the substrate directs the orientation of the Fe(IV)=O intermediate, presenting specific frontier molecular orbitals that can activate either selective halogenation or hydroxylation.

SyrB2, an α -ketoglutarate (α KG)-dependent NHF enzyme found in *Pseudomonas syringae* pv. *syringae*, halogenates the methyl group of L-threonine (L-Thr) using non-ribosomal peptide synthetase machinery¹⁸. The Fe(II) active site is ligated by two histidines and one halide (Cl[−] or Br[−]) (see Supplementary Fig. 1)¹⁹, in contrast to the two-histidine/one-carboxylate ‘facial triad’ of other NHF enzymes^{1,2,4,5,8}. Although the mechanisms of O₂ activation leading to the highly-reactive $S = 2$ Fe(IV)=O intermediate are thought to be similar for all α KG-dependent NHF enzymes, there is a notable divergence in their subsequent catalytic cycles (Fig. 1). The Fe(IV)=O species abstracts a hydrogen atom from the substrate to form an Fe(III)–OH species and a substrate radical; in hydroxylases, the subsequent step is HO[•] rebound to form a hydroxylated product^{1,2}, but in SyrB2 the native L-Thr substrate is chlorinated instead, and the 4-Cl-L-Thr product is used in the biosynthesis of the phytotoxin syringomycin E¹⁸. Owing to their reactivity, Fe(IV)=O intermediates in enzyme reactions are challenging to trap and characterize. For SyrB2, however, use of the non-native substrate L-cyclopropylglycine (L-Cpg) and the heterologous substrate carrier protein CytC2 has provided a long-lived species at the concentrations required for spectroscopic investigation^{17,20}.

Nuclear resonance vibrational spectroscopy (NRVS) utilizes third-generation synchrotron radiation to probe the vibrational sidebands of the ⁵⁷Fe Mössbauer nuclear resonance peak at 14.4 keV (refs 21–23). NRVS is a site-selective technique allowing the observation of only normal modes involving Fe motion, which makes it ideal for studying iron-dependent enzymes without interference from protein backbone modes. (SyrB2)Fe(IV)=O can be generated in high purity with both Cl[−] and Br[−] ligation of the Fe(IV)=O unit, providing a mass perturbation that aids in the assignment of NRVS peaks and ultimately the structure of the intermediate. The NRVS methodology is coupled with spectroscopically calibrated density functional theory (DFT) calculations to evaluate specific frontier molecular orbitals (FMOs) responsible for hydrogen-atom abstraction that can selectively lead to halogenation or hydroxylation, depending on the substrate.

In Fig. 2 we show the NRVS partial vibrational density-of-states (PVDOS) spectra of (SyrB2)Cl–Fe(IV)=O with L-Cpg–S–CytC2 bound (referred to as SyrB2–Cl) and the bromide analogue, (SyrB2)Br–Fe(IV)=O with L-Cpg–S–CytC2 bound (SyrB2–Br). For practical reasons (see Methods), data collection was restricted to <600 cm^{−1}; modes in this region are affected by large (Cl versus Br; see below) but not small (¹⁶O versus ¹⁸O; see Supplementary Fig. 2) mass perturbations. There are three distinct features for each species, as indicated by the bracketed energy regions: region 1 (340–400 cm^{−1}), region 2 (285–340 cm^{−1}) and region 3 (200–285 cm^{−1}). For the higher-energy regions 1 and 2, the

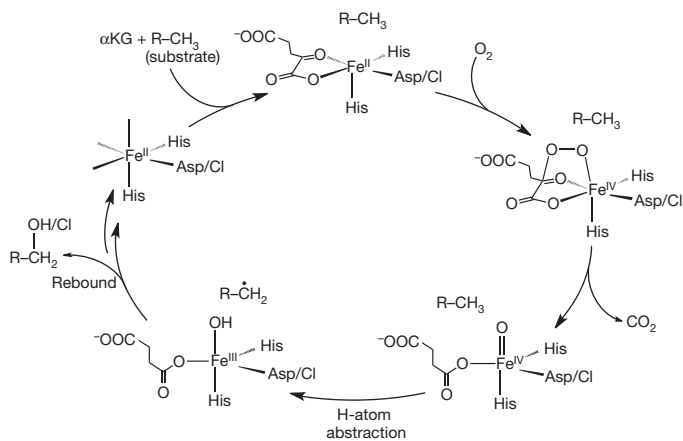


Figure 1 | Catalytic cycle of α KG-dependent NHF enzymes. α KG and substrate binding induces a six-coordinate to five-coordinate conversion (top), providing a site for O₂ to bind and form an Fe(IV)–peroxo species that nucleophilically attacks α KG, producing a peroxo-bridged Fe(IV) species (right)²⁴. Decarboxylation of α KG leads to the reactive Fe(IV)=O intermediate (bottom right), which goes on to perform hydrogen-atom abstraction (bottom left) and subsequent rebound hydroxylation or halogenation.

¹Department of Chemistry, Stanford University, Stanford, California 94305, USA. ²Department of Chemistry, Pennsylvania State University, University Park, Pennsylvania 16802, USA. ³Argonne National Laboratory, APS/XFD, 431/D003, Argonne, Illinois 60439, USA. ⁴Spring-8, JASRI, Hyogo 679-5198, Japan. ⁵Research Reactor Institute, Kyoto University, Osaka 590-0494, Japan. ⁶Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, Pennsylvania 16802, USA. ⁷SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [†]Present address: Department of Chemical Physiology, The Scripps Research Institute, La Jolla, California 92037, USA.

*These authors contributed equally to this work.

peaks of SyrB2-Cl are more intense. However, for the low-energy region 3, the peak envelope for SyrB2-Br is considerably more intense and shifted to lower energy.

Previous computational studies of the Cl-Fe(IV)=O intermediate of SyrB2 predicted six-coordinate structures with the succinate bound as a bidentate ligand to Fe (refs 9–12). The DFT-calculated NRVs spectra of these six-coordinate structures (Supplementary Fig. 3) do not reproduce the splitting pattern and intensity distribution of the experimental data, and can thus be eliminated from consideration.

To generate and evaluate suitable structural candidates, the O₂ reaction coordinate taking SyrB2 to its Fe(IV)=O intermediate (Fig. 1) was investigated using DFT calculations. The initial structure was taken from the crystal structure of the SyrB2 Fe(II) active site with the α KG cofactor and Cl⁻ bound (Supplementary Fig. 1)¹⁹ and the native substrate L-Thr positioned according to a molecular docking procedure¹²; its side chain was also modified into the non-native substrate L-Cpg to generate a second starting structure. Application of the spectroscopically calibrated DFT methodology used for a related α KG-dependent mononuclear NHE enzyme²⁴ resulted in an equivalent O₂ reaction coordinate for the Fe(II) active site of SyrB2 (Supplementary Figs 4 and 5a, b).

This O₂ reaction coordinate leads to **1**_{Cpg}-Cl (Fig. 3b) and **1**_{Thr}-Cl (Supplementary Fig. 5b) with L-Cpg and L-Thr respectively; both are five-coordinate trigonal bipyramidal (TBP) Fe(IV)=O structures possessing an axial oxo group and a monodentate succinate. Significantly, in both cases, the Fe-oxo vector is oriented perpendicularly to the target substrate C-H bond, with interesting implications for π -channel reactivity^{25,26}.

1_{Cpg}-Cl was evaluated as a structural candidate for the Fe(IV)=O intermediate in the NRVs sample. The geometry-optimized Fe-oxo and Fe-Cl bond lengths (Fig. 3b) of **1**_{Cpg}-Cl are in close agreement with the experimental extended X-ray absorption fine structure (EXAFS) values (1.66 Å and 2.31 Å respectively)⁷. The Br⁻ cognate, **1**_{Cpg}-Br, was generated by replacing Cl⁻ with Br⁻ and reoptimizing the structure; its Fe-Br bond length of 2.45 Å (Fig. 3b) agrees well with the EXAFS value of 2.43 Å for the related halogenase CytC3 (ref. 6). Thus, these five-coordinate TBP intermediates **1**_{Cpg}-X (X = Cl or Br) resulting from the O₂ reaction coordinate were used for comparison with the experimental NRVs data on the SyrB2 Fe(IV)=O intermediates SyrB2-X.

As seen in Fig. 3, the five-coordinate TBP species **1**_{Cpg}-X result in DFT-predicted spectra that reproduce the experimental spectra. First,

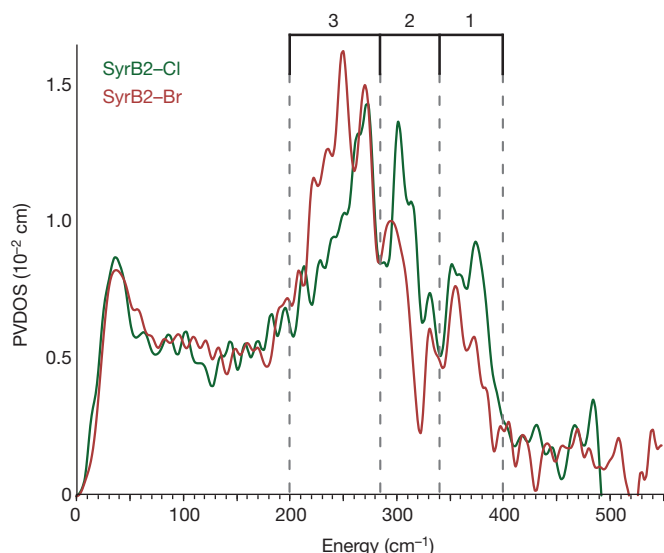


Figure 2 | NRVs PVDOS spectra of SyrB2-Cl and SyrB2-Br. Regions 3, 2 and 1 that contain intense features are indicated at the top.

there are three distinct peaks falling within the energy regions of 200–285 cm⁻¹, 285–340 cm⁻¹ and 340–400 cm⁻¹, matching regions 3, 2 and 1 in Fig. 2. Second, the intensities of the peaks in the two higher-energy regions are greater for **1**_{Cpg}-Cl than for **1**_{Cpg}-Br, and the intensity of the peak envelope in the lowest-energy region for **1**_{Cpg}-Br is greater and shifted to lower energy with respect to that of **1**_{Cpg}-Cl, reproducing the spectral intensity distributions of the experimental data (Fig. 2). Other five- and six-coordinate structures were generated as possible candidates for the Fe(IV)=O species, starting from **1**_{Cpg}-X and shifting either the Fe-ligating atoms or the hydrogen-bonding network to the oxo group (Supplementary Fig. 6). From the predicted NRVs spectra of these structures and of the structures generated in previous (computational) studies (Supplementary Fig. 3), all structures except five-coordinate TBP can be eliminated owing to their poor agreement with the experimental NRVs data.

Correlating the DFT-calculated spectra of five-coordinate TBP **1**_{Cpg}-X (Fig. 3) with the experimental spectra (Fig. 2), the NRVs peaks can be assigned to four normal modes (Fig. 4): the feature in region 1 (Fig. 2) originates from the Fe-succinate stretch; the feature in region 2 is composed of a pair of *trans*-axial bending modes (these would be degenerate in strict TBP symmetry, but are calculated to split in energy because of the wider equatorial X-Fe-succinate angle of 143°)¹⁵; the

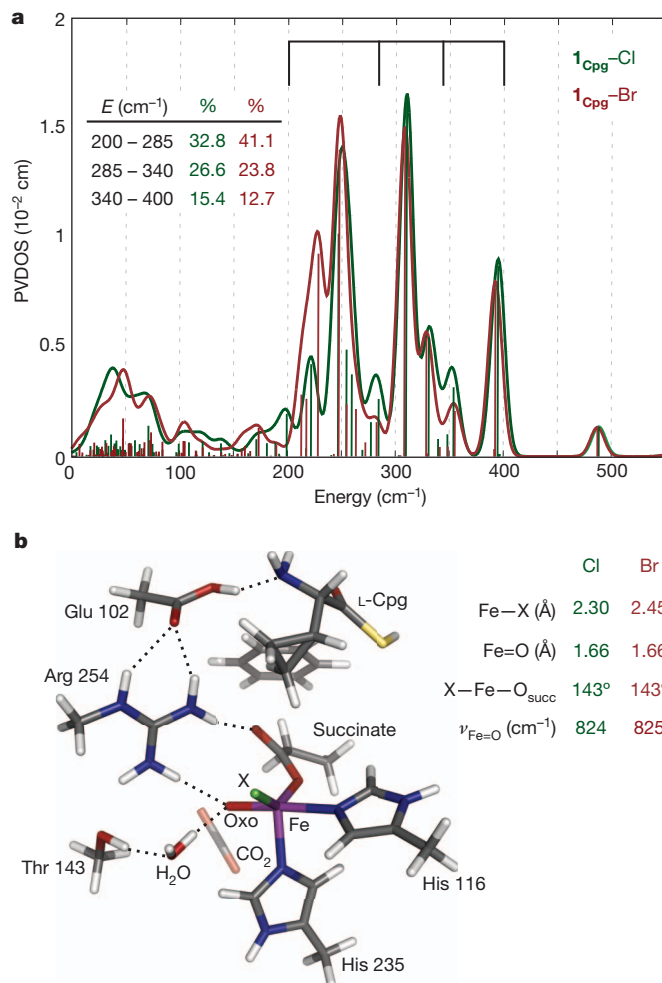


Figure 3 | Computational spectra and structure of five-coordinate TBP structural candidate **1**_{Cpg}-X for the Fe(IV)=O intermediate of SyrB2. **a**, DFT-predicted PVDOS NRVs spectra. Vertical bars represent relative calculated mode-composition factors of vibrational modes, and brackets at top correspond to energy regions from Fig. 2. Inset, peak intensity contributions (from the three bracketed regions) to the overall PVDOS envelope. **b**, Structure of **1**_{Cpg}-X (left), along with geometric parameters and Fe-oxo stretching frequencies (right).

lowest-energy region 3 has a peak envelope calculated to contain the *trans*-axial stretch and the Fe–X stretch, with the Fe–Br stretch being lower in energy by 30 cm^{−1} and more intense by 1.5 times. The redistribution in intensity is attributed to the mass perturbation of the Br, which has almost no motion in the Fe–X stretching mode and consequently induces greater Fe motion in the mode. This Fe motion is borrowed from higher-energy modes, as analysed in Supplementary Fig. 7. The NRVs peak pattern of SyrB2–X parallels that of a crystallographically characterized TBP *S* = 2 Fe(IV)=O model complex (Supplementary Fig. 8)¹⁵, further demonstrating the sensitivity of NRVs to geometric structure.

We note that two distinct Fe(IV) species are detected by Mössbauer spectroscopy, differing in quadrupole splitting, ΔE_Q (Supplementary Fig. 9), in each Fe(IV)=O intermediate generated⁷. A possible explanation for this speciation lies in the hydrogen-bonding interactions with the oxo group: **1**_{Cpg}–Cl has two (with Arg 254 and H₂O), whereas **1**_{Thr}–Cl has one (with H₂O). Their predicted NRVs spectra are similar (Supplementary Figs 6 and 10a), but their calculated values of ΔE_Q are different, with that of **1**_{Cpg}–Cl (−0.50 mm s^{−1}) being smaller in magnitude than that of **1**_{Thr}–Cl (−0.71 mm s^{−1}). Decreasing the number of hydrogen bonds strengthens the Fe–oxo bond, thus increasing the magnitude of (negative) ΔE_Q (Supplementary Fig. 10a and Supplementary Table 1). These calculations suggest that variability in hydrogen-bonding interactions with the oxo group results in Fe(IV) speciation, not some structural difference.

For the native L-Thr substrate (starting from the O₂-reaction-coordinate-derived Fe(IV)=O species **1**_{Thr}–Cl, which has its Fe–oxo vector perpendicular to C–H (Supplementary Fig. 5b)) the hydrogen-atom abstraction reactivity was computationally evaluated (Fig. 5). The target C–H approaches in a π -trajectory, transferring an α -electron into the oxo π -FMO, resulting in an Fe(III)–OH first product (**1**_{Thr}–Fe(III)–OH), in which the Fe(III) has spin *S* = 5/2. The free-energy barrier ΔG^\ddagger for this π -pathway is +100.4 kJ mol^{−1}, in reasonable agreement with the experimental value of +79.4 kJ mol^{−1} (ref. 7). A number of possible explanations have been considered for the subsequent Cl[•] rebound^{17,27}; here we show that in this first product, the substrate radical is positioned closer to the Cl than to the OH ligand of Fe(III) (consideration of their ionic radii places Cl 0.5 Å closer than OH), and OH is also stabilized by hydrogen-bonding to succinate (Fig. 5, right). This conformation disfavours HO[•] rebound but is well oriented for Cl[•] rebound, as observed experimentally with the native substrate. This perpendicular Fe–oxo orientation is, for the six-coordinate structures proposed in previous computational studies^{9–12}, inaccessible via the O₂-activation pathway because a bidentate succinate would block the oxo group from reorienting.

Evaluation of the reaction coordinate for O–O cleavage leading to the Fe(IV)=O species (Supplementary Fig. 5) revealed that the positioning of L-Thr is fixed by two hydrogen-bonding interactions (–OH and –NH₃⁺) to Glu 102. This configuration results in the perpendicular orientation of the Fe(IV)–oxo vector relative to the substrate C–H bond. However, the alternative substrate L-norvaline (L-Nva) lacks the –OH group, and thus its –NH₃⁺ group can rotate to form a hydrogen bond with the O–O (peroxy) bridge, leading to a structure with an

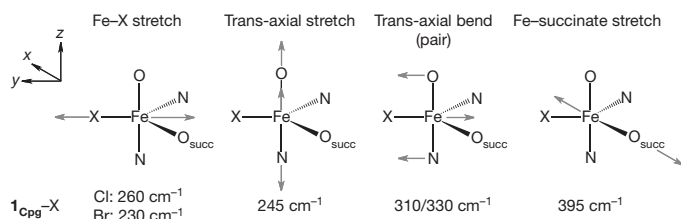


Figure 4 | DFT-predicted normal modes of **1**_{Cpg}–X. Four normal modes of the five-coordinate TBP Fe(IV)=O structure are shown; above are the mode descriptions, below are the corresponding frequencies. The Fe–oxo vector defines the z axis.

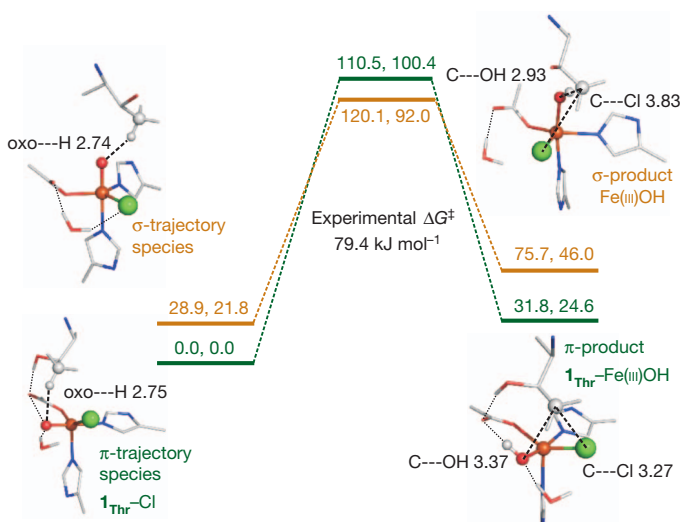


Figure 5 | Hydrogen-atom abstraction reaction coordinates. Energies (given as ΔE , ΔG in kJ mol^{−1}) for the π -trajectory (**1**_{Thr}–Cl, green numerals and levels) and the σ -trajectory (orange numerals and levels) are shown. Reactants are on the left; energies for the transition state (where the Fe has developed *S* = 5/2 spin character) are given at the top; structures of Fe(III)–OH products (in which the Fe(III) has spin *S* = 5/2) are displayed on the right, showing the π -product **1**_{Thr}–Fe(III)–OH (with hydrogen-bonding interactions indicated) set up for chlorination and the σ -product set up for hydroxylation. Distances are given in Å. For additional structural details, see Supplementary Fig. 11.

Fe–oxo vector oriented towards the substrate C–H bond (Supplementary Fig. 5c). An analogous L-Thr orientation was thus generated to evaluate its hydrogen-atom abstraction trajectory while maintaining the same C–H bond (Fig. 5, left).

For this orientation, the C–H approaches the Fe–oxo unit in a σ -trajectory, transferring an α -electron into the oxo σ -FMO to give an Fe(III)–OH product in which the Fe(III) has spin *S* = 5/2; this σ -pathway has a ΔG^\ddagger of +70.2 kJ mol^{−1} (Fig. 5). Relative to the π -pathway Fe(III)–OH product, this Fe(III)–OH has the substrate radical closer to the OH ligand than the Cl (by 0.5 Å, based on ionic radii), and OH has no hydrogen-bonding partner. The parallel Fe–oxo orientation therefore favours HO[•] rebound, as is observed experimentally for the non-native substrate L-Nva¹⁷. We also note that the barrier for π -attack is somewhat higher than that for σ -attack, which is consistent with the higher barrier observed experimentally for halogenation relative to hydroxylation (by ~17 kJ mol^{−1})¹⁷, reflecting halogenation selectivity over efficiency.

We have used NRVs to characterize the NHFe enzyme oxo intermediate in SyrB2 and found the Fe to be in a five-coordinate TBP configuration, with an axial Fe(IV)=O bond. The native-substrate-bound O₂ reaction coordinate reproduces this structure and gives an intermediate with its Fe–oxo vector perpendicular to the substrate C–H bond; this Fe–oxo orientation is active in hydrogen-atom abstraction, via its π^* -FMO. This positions the substrate radical favourably for Cl[•] rebound, thus defining a selective mechanism in halogenases for chlorination of the native substrate. Alternatively, with a non-native substrate, variation in the O₂ reaction coordinate can lead to an intermediate with its Fe–oxo vector parallel to the substrate C–H bond, leading to hydrogen-atom abstraction via a σ -pathway and a substrate radical positioned for HO[•] rebound and resultant hydroxylation.

METHODS SUMMARY

Sample preparation. Samples were prepared essentially as previously described⁷. Details of experimental conditions (such as concentrations) are provided in Supplementary Fig. 9.

Spectroscopic methods. ⁵⁷Fe NRVs spectra were recorded on multiple occasions at Beamline 3-ID at the Advanced Photon Source of the Argonne National Laboratory²⁸, at temperatures between 10 K and 30 K, and once at BL09XU at

SPring-8 in Japan²⁹, at temperatures between 50 K and 80 K. NRVS data were collected only up to 600 cm⁻¹ owing to relatively low sample concentrations (~1.8 mM); because NRVS signal intensity decreases exponentially with increasing energy, data collected above this region (given our allocated beamtime) would have had a low signal-to-noise ratio.

Computational methods. DFT methods were used to simulate the NRVS spectra and Mössbauer parameters of all structural candidates, and to evaluate the hydrogen-atom abstraction reaction coordinates (see Methods).

Full Methods and any associated references are available in the online version of the paper.

Received 4 February; accepted 16 May 2013.

- Solomon, E. I. *et al.* Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **100**, 235–350 (2000).
- Costas, M., Mehn, M. P., Jensen, M. P. & Que, L. Jr. Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem. Rev.* **104**, 939–986 (2004).
- Vaillancourt, F. H., Yeh, E., Vosburg, D. A., Garneau-Tsodikova, S. & Walsh, C. T. Nature's inventory of halogenation catalysts: oxidative strategies predominate. *Chem. Rev.* **106**, 3364–3378 (2006).
- Krebs, C., Galonić Fujimori, D., Walsh, C. T. & Bollinger, J. M. Jr. Non-heme Fe(IV)-oxo intermediates. *Acc. Chem. Res.* **40**, 484–492 (2007).
- Eser, B. E. *et al.* Direct spectroscopic evidence for a high-spin Fe(IV) intermediate in tyrosine hydroxylase. *J. Am. Chem. Soc.* **129**, 11334–11335 (2007).
- Galonić Fujimori, D. *et al.* Spectroscopic evidence for a high-spin Br-Fe(IV)-oxo intermediate in the alpha-ketoglutarate-dependent halogenase CytC3 from *Streptomyces*. *J. Am. Chem. Soc.* **129**, 13408–13409 (2007).
- Matthews, M. L. *et al.* Substrate-triggered formation and remarkable stability of the C–H bond-cleaving chloroferryl intermediate in the aliphatic halogenase, SyrB2. *Biochemistry* **48**, 4331–4343 (2009).
- Panay, A. J., Lee, M., Krebs, C., Bollinger, J. M. Jr & Fitzpatrick, P. F. Evidence for a high-spin Fe(IV) species in the catalytic cycle of a bacterial phenylalanine hydroxylase. *Biochemistry* **50**, 1928–1933 (2011).
- Pandian, S., Vincent, M. A., Hillier, I. H. & Burton, N. A. Why does the enzyme SyrB2 chlorinate, but does not hydroxylate, saturated hydrocarbons? A density functional theory (DFT) study. *Dalton Trans.* 6201–6207 (2009).
- Kulik, H. J., Blasiak, L. C., Marzari, N. & Drennan, C. L. First-principles study of non-heme Fe(II) halogenase SyrB2 reactivity. *J. Am. Chem. Soc.* **131**, 14426–14433 (2009).
- de Visser, S. P. & Latifi, R. Carbon dioxide: a waste product in the catalytic cycle of alpha-ketoglutarate dependent halogenases prevents the formation of hydroxylated by-products. *J. Phys. Chem. B* **113**, 12–14 (2009).
- Borowski, T., Noack, H., Radoń, M., Zych, K. & Siegbahn, P. E. M. Mechanism of selective halogenation by SyrB2: a computational study. *J. Am. Chem. Soc.* **132**, 12887–12898 (2010).
- Usharani, D., Janardanan, D. & Shaik, S. Does the TauD enzyme always hydroxylate alkanes, while an analogous synthetic non-heme reagent always desaturates them? *J. Am. Chem. Soc.* **133**, 176–179 (2011).
- Bell, C. B. *et al.* A combined NRVS and DFT study of Fe^{IV}=O model complexes: a diagnostic method for the elucidation of non-heme iron enzyme intermediates. *Angew. Chem. Int. Edn* **47**, 9071–9074 (2008).
- Wong, S. D. *et al.* Nuclear resonance vibrational spectroscopy on the Fe^{IV}=O S = 2 non-heme site in TMG₃tren: experimentally calibrated insights into reactivity. *Angew. Chem. Int. Edn* **50**, 3215–3218 (2011).
- Park, K. *et al.* Nuclear resonance vibrational spectroscopic and computational study of high-valent diiron complexes relevant to enzyme intermediates. *Proc. Natl Acad. Sci. USA* **110**, 6275–6280 (2013).
- Matthews, M. L. *et al.* Substrate positioning controls the partition between halogenation and hydroxylation in the aliphatic halogenase, SyrB2. *Proc. Natl Acad. Sci. USA* **106**, 17723–17728 (2009).
- Vaillancourt, F. H., Yin, J. & Walsh, C. T. SyrB2 in syringomycin E biosynthesis is a nonheme Fe^{II} α-ketoglutarate- and O₂-dependent halogenase. *Proc. Natl Acad. Sci. USA* **102**, 10111–10116 (2005).
- Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T. & Drennan, C. L. Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440**, 368–371 (2006).
- Krebs, C. *et al.* Novel approaches for the accumulation of oxygenated intermediates to multi-millimolar concentrations. *Coord. Chem. Rev.* **257**, 234–243 (2013).
- Seto, M., Yoda, Y., Kikuta, S., Zhang, X. & Ando, M. Observation of nuclear resonant scattering accompanied by phonon excitation using synchrotron radiation. *Phys. Rev. Lett.* **74**, 3828–3831 (1995).
- Chumakov, A. I. & Sturhahn, W. Experimental aspects of inelastic nuclear resonance scattering. *Hyperfine Interact.* **123/124**, 781–808 (1999).
- Sage, J. T. *et al.* Nuclear resonance vibrational spectroscopy of a protein active-site mimic. *J. Phys. Condens. Matter* **13**, 7707–7722 (2001).
- Diebold, A. R. *et al.* Activation of α-keto acid-dependent dioxygenases: application of an {FeNO}⁷/FeO₂⁸ methodology for characterizing the initial steps of O₂ activation. *J. Am. Chem. Soc.* **133**, 18148–18160 (2011).
- Neidig, M. L. *et al.* Spectroscopic and electronic structure studies of aromatic electrophilic attack and hydrogen-atom abstraction by non-heme iron enzymes. *Proc. Natl Acad. Sci. USA* **103**, 12966–12973 (2006).
- Srnc, M., Wong, S. D., England, J., Que, L. & Solomon, E. I. π-frontier molecular orbitals in S = 2 ferryl species and elucidation of their contributions to reactivity. *Proc. Natl Acad. Sci. USA* **109**, 14326–14331 (2012).
- Comba, P. & Wunderlich, S. Iron-catalyzed halogenation of alkanes: modeling of nonheme halogenases by experiment and DFT calculations. *Chemistry* **16**, 7293–7299 (2010).
- Alp, E. E., Mooney, T. M., Toellner, T. & Sturhahn, W. Nuclear resonant scattering beamline at the Advanced Photon Source. *Hyperfine Interact.* **90**, 323–334 (1994).
- Yoda, Y. *et al.* Nuclear resonant scattering beamline at SPring-8. *Nucl. Instrum. Methods A* **467–468**, 715–718 (2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements Funding for this work was provided by the National Institutes of Health (GM-40392 to E.I.S. and GM-69657 to J.M.B. and C.K.) and the National Science Foundation (MCB-0919027 to E.I.S., and MCB-642058 and CHE-724084 to J.M.B. and C.K.). Work at the Advanced Photon Source was supported by the Department of Energy, Office of Science, under contract DE-AC-02-06CH11357. Synchrotron experiments at SPring-8 were performed with the approval of the Japan Synchrotron Radiation Research Institute (JASRI; proposal no. 2010B1569). M.S. thanks the Rulišek group at the IOCB, Prague, for use of their computational resources.

Author Contributions S.D.W. and M. Srnc contributed equally to this work. E.I.S., C.K. and J.M.B. designed the experiments. S.D.W., M. Srnc, M.L.M., L.V.L., Y.K., K.P. and C.B.B. performed the experiments. S.D.W., M. Srnc and E.I.S. analysed the data and wrote the manuscript. E.E.A., J.Z., Y.Y., S.K. and M. Seto provided technical assistance at the synchrotron beamlines.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.K. (ckrebs@psu.edu), J.M.B. (jmb21@psu.edu) or E.I.S. (edward.solomon@stanford.edu).

METHODS

Sample preparation. Samples were prepared essentially as previously described⁷. Details of experimental conditions (such as concentrations) are provided in Supplementary Fig. 9.

Spectroscopic methods. ⁵⁷Fe NRVs spectra were recorded on multiple occasions at Beamline 3-ID at the Advanced Photon Source of the Argonne National Laboratory²⁸, at temperatures between 10 K and 30 K, and once at BL09XU at SPring-8 in Japan²⁹, at temperatures between 50 K and 80 K. NRVs data were collected only up to 600 cm⁻¹ owing to relatively low sample concentrations (~1.8 mM); because NRVs signal intensity decreases exponentially with increasing energy, data collected above this region (given our allocated beamtime) would have had a low signal-to-noise ratio.

Computational methods. DFT methods were used to simulate the NRVs spectra and Mössbauer parameters of all structural candidates, and to evaluate the hydrogen-atom abstraction reaction coordinates. Spin-unrestricted DFT calculations were performed using the Turbomole 6.3³⁰ and Gaussian 09³¹ programs. Turbomole 6.3 was used to perform geometry optimizations and frequency calculations of the structural candidates in Supplementary Fig. 6, with the BP86^{32–34} exchange-correlation functional and the double- ζ def2-SVP basis set³⁵. Single-point energies were recomputed using the larger triple- ζ basis set def2-TZVP³⁵. Turbomole calculations were expedited by expanding the Coulomb integrals in an auxiliary basis set, using the RI-J approximation^{36,37}. Solvation effects were taken into account by using the conductor-like screening model (COSMO) method^{38,39} with a dielectric constant $\epsilon_r = 4$ as is appropriate for the protein environment (the COSMO radii were set as follows: H, 1.30 Å; C, 2.00 Å; N, 1.83 Å; O, 1.72 Å; Cl, 2.05 Å; Br, 2.16 Å; S, 2.16 Å; and Fe, 2.23 Å). This is referred to as the RI-BP86/def2-SVP (or def2-TZVP)/COSMO approach or level of theory.

Gaussian 09 was used to perform geometry optimizations and frequency calculations of the structural candidates **1**Cpg-X in Fig. 3, with the functional/basis set combination BP86/6-311G*^{40–43}. Solvation effects were taken into account with the polarized continuum model (PCM)^{44–47}, using $\epsilon = 4.0$. This is referred to as the BP86/6-311G*/PCM approach or level of theory.

NRVS PVDOS spectra were simulated by fitting the DFT-calculated mode composition factor⁴⁸

$$e_{n,\text{Fe}}^2 = \frac{m_{\text{Fe}} r_{n,\text{Fe}}^2}{\sum_i m_i r_{n,i}^2}$$

(where m is the mass of atom i , and r is the displacement of atom i) for each normal mode n with individual Gaussians of FWHM 15 cm⁻¹, using the *gennrvs* script⁴⁹.

Mössbauer isomer shifts and quadrupole splittings were calculated according to published methods⁵⁰.

The initial structure of the SyrB2 Fe(II) active site used for DFT calculations was taken from its crystal structure (Supplementary Fig. 1 and ref. 19). Inclusion of the substrate (L-Cpg-SH, where the terminal -SH group represents truncation at the thioester linkage to the phosphopantetheine cofactor) was modelled according to ref. 12. Except where stated, the S atom of the substrate was frozen during geometry optimization. Note that the substrate does indeed fit well in the cavity of the active site (Supplementary Fig. 1d).

The O₂-reaction coordinate starting from the SyrB2-Cl Fe(II) active site was pursued analogously to ref. 24, at the RI-BP86/def2-SVP(def2-TZVP)/COSMO level of theory. The complete O₂-reaction coordinate for the L-Cpg-bound active site (with either Cl⁻ or Br⁻) is shown in Supplementary Fig. 4, and Supplementary Fig. 5 shows the final O–O cleavage step, leading to the Fe(IV)=O intermediate, for three versions of the active site containing L-Cpg (inert substrate), L-Thr (native substrate) and L-Nva (non-native substrate).

The hydrogen-atom abstraction reaction coordinates of the SyrB2-Cl Fe(IV)=O intermediate were evaluated using the Turbomole 6.3 program³⁰. **1**Thr-Cl was optimized at the B3LYP^{51–53}+D2/def2-SVP level (where +D2 stands for the second version of Grimme's empirical dispersion correction^{54,55}). Thermodynamic corrections to give enthalpic (ΔH) and Gibbs (ΔG) energies were calculated at $T = 278.15$ K to reproduce experimental conditions⁷. Single-point energies were calculated at the B3LYP+D2/def2-TZVP/ COSMO($\epsilon_r = 4.0$) level. The calculated NRVs spectra of these Thr-bound species (Supplementary Fig. 10) are similar to those of their L-Cpg-bound counterparts (Supplementary Fig. 6),

showing that the substrate does not affect the NRVs spectra because it is not directly coordinated to the Fe centre. Starting with **1**Thr-Cl and **2**Thr-Cl as the reactant complexes, each hydrogen-atom abstraction reaction was pursued along the oxo-H(L-Thr) coordinate and each transition state was optimized from the highest-energy structure along the reaction coordinate. An internal reaction coordinate was calculated from each optimized transition state (forward) to obtain the product (Fe(III)-OH + substrate radical) and (backward) to confirm the validity of the reactant complex structure.

- Ahlrichs, R., Bär, M., Häser, M., Horn, H. & Kölmel, C. Electronic structure calculations on workstation computers: the program system Turbomole. *Chem. Phys. Lett.* **162**, 165–169 (1989).
- Frisch, M. J. et al. Gaussian 09, Revision A.1 (Gaussian, Wallingford, 2009).
- Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).
- Perdew, J. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **33**, 8822–8824 (1986).
- Vosko, S. H., Wilk, L. & Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **58**, 1200–1211 (1980).
- Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).
- Eichkorn, K., Treutler, O., Öhm, H., Häser, M. & Ahlrichs, R. Auxiliary basis sets to approximate Coulomb potentials. *Chem. Phys. Lett.* **240**, 283–290 (1995).
- Eichkorn, K., Weigend, F., Treutler, O. & Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chim. Acta* **97**, 119–124 (1997).
- Klamt, A. & Schuurmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perkin Trans. 2* **0**, 799–805 (1993).
- Schäfer, A., Klamt, A., Sattel, D., Lohrenz, J. C. W. & Eckert, F. COSMO implementation in Turbomole: extension of an efficient quantum chemical code towards liquid systems. *Phys. Chem. Chem. Phys.* **2**, 2187–2193 (2000).
- Wachters, A. Gaussian basis set for molecular wavefunctions containing third-row atoms. *J. Chem. Phys.* **52**, 1033–1036 (1970).
- Hay, P. J. Gaussian basis sets for molecular calculations. The representation of 3d orbitals in transition-metal atoms. *J. Chem. Phys.* **66**, 4377–4384 (1977).
- McLean, A. D. & Chandler, G. S. Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, $Z = 11$ –18. *J. Chem. Phys.* **72**, 5639–5648 (1980).
- Krishnan, R., Binkley, J. S., Seeger, R. & Pople, J. A. Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J. Chem. Phys.* **72**, 650–654 (1980).
- Mennucci, B. & Tomasi, J. Continuum solvation models: a new approach to the problem of solute's charge distribution and cavity boundaries. *J. Chem. Phys.* **106**, 5151–5158 (1997).
- Mennucci, B., Cancès, E. & Tomasi, J. Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: theoretical bases, computational implementation, and numerical applications. *J. Phys. Chem. B* **101**, 10506–10517 (1997).
- Cammi, R., Mennucci, B. & Tomasi, J. Second-order Møller–Plesset analytical derivatives for the polarizable continuum model using the relaxed density approach. *J. Phys. Chem. A* **103**, 9100–9108 (1999).
- Cammi, R., Mennucci, B. & Tomasi, J. Fast evaluation of geometries and properties of excited molecules in solution: a Tamm-Dancoff model with application to 4-dimethylamino-enzonitrile. *J. Phys. Chem. A* **104**, 5631–5637 (2000).
- Leu, B. M. et al. Quantitative vibrational dynamics of iron in nitrosyl porphyrins. *J. Am. Chem. Soc.* **126**, 4211–4227 (2004).
- Tenderholt, A. *gennrvs* (2009); Pymol script available at <http://www.stanford.edu/group/solomon/gennrvs/gennrvs.py.txt>.
- Srncic, M. et al. Structural and spectroscopic properties of the peroxodiferic intermediate of *Ricinus communis* soluble Δ^9 desaturase. *Inorg. Chem.* **51**, 2806–2820 (2012).
- Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
- Lee, C., Yang, W. & Parr, R. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
- Miehlich, B., Savin, A., Stoll, H. & Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem. Phys. Lett.* **157**, 200–206 (1989).
- Grimme, S. Accurate description of van der Waals complexes by density functional theory including empirical corrections. *J. Comput. Chem.* **25**, 1463–1473 (2004).
- Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).

Exciting Andreev pairs in a superconducting atomic contact

L. Bretheau^{1*}, Ç. Ö. Girit^{1*}, H. Pothier¹, D. Esteve¹ & C. Urbina¹

The Josephson effect describes the flow of supercurrent in a weak link—such as a tunnel junction, nanowire or molecule—between two superconductors¹. It is the basis for a variety of circuits and devices, with applications ranging from medicine² to quantum information³. Experiments using Josephson circuits that behave like artificial atoms⁴ are now revolutionizing the way we probe and exploit the laws of quantum physics^{5,6}. Microscopically, the supercurrent is carried by Andreev pair states, which are localized at the weak link. These states come in doublets and have energies inside the superconducting gap^{7–10}. Existing Josephson circuits are based on properties of just the ground state of each doublet, and so far the excited states have not been directly detected. Here we establish their existence through spectroscopic measurements of superconducting atomic contacts. The spectra, which depend on the atomic configuration and on the phase difference between the superconductors, are in complete agreement with theory. Andreev doublets could be exploited to encode information in novel types of superconducting qubits^{11–13}.

A bulk, isolated Bardeen–Cooper–Schrieffer superconductor can be described by a density of states (DOS) having a gap around the Fermi energy of 2Δ , which is the minimum energy necessary to excite an electron pair¹⁴. In the presence of a short weak link, the superconducting phase can be easily twisted, leading to a local modification of the DOS and the creation of new states inside the gap. These Andreev bound states have energies $\pm E_A$, with E_A given by

$$E_A = \Delta \sqrt{1 - \tau \sin^2(\delta/2)} \quad (1)$$

for a weak link which has a phase difference δ across it and a single conduction channel of transmission probability τ (Fig. 1a). As long as $E_A < \Delta$, these states cannot propagate into the bulk superconductor and are bound to the weak link, over a distance of order ξ , the superconducting coherence length. The ground Andreev pair state $|-\rangle$ has energy $-E_A$, and the lowest possible pair excitation of the system, requiring an energy $2E_A$, is a transition to the excited Andreev pair state $|+\rangle$ at $+E_A$. The phase dependence of $\pm E_A$ gives rise to opposite supercurrents for the two states, $\mp (2\pi/\phi_0)(\partial E_A/\partial \delta)$, with $\phi_0 = h/2e$ the flux quantum (h is Planck's constant, and e the charge of an electron).

Current Josephson circuits are primarily based on tunnel Josephson junctions, which have many conduction channels with small transmission probabilities ($\tau \ll 1$). In this limit, the ground state energy $-E_A$ in each channel is proportional to $-\cos\delta$. Summing over all channels, one recovers the standard Josephson coupling energy $-E_J \cos\delta$ and the sinusoidal current–phase relation predicted by Josephson¹. For channels of arbitrary transmissions, the ground state $|-\rangle$ has been probed through measurements of the current–phase relation in superconducting atomic contacts¹⁵. Excitations created by the addition or removal of an electron from the state $|-\rangle$ have been observed in superconducting atomic contacts¹⁶ and quantum dots connected to superconductors^{17,18}. The continuum of Andreev states that form in superconductor–normal–superconductor structures has also been probed^{19,20}. Thermal occupation of the excited states was invoked to explain the temperature

dependence of the supercurrent¹⁰. However, the excited Andreev pair state $|+\rangle$ has not been directly detected. Here we present spectroscopic evidence of excited Andreev pair states in superconducting atomic

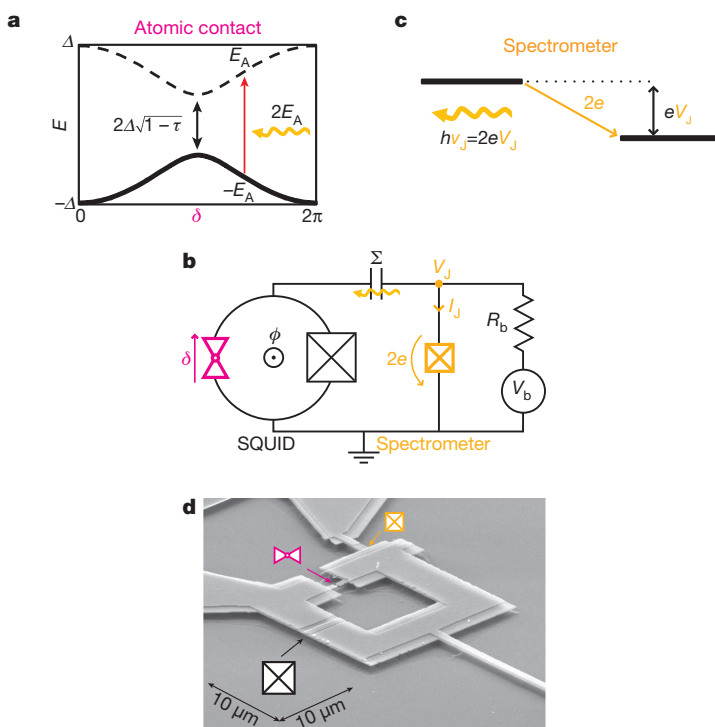


Figure 1 | Principles of spectroscopy of the Andreev transition. **a**, Phase (δ) dependence of the Andreev levels with energies $\pm E_A$ in a short transport channel of transmission τ . Energy Δ is the superconducting gap. Andreev transitions from the ground state to the excited state (red arrow) are induced by photons (orange wavy arrow). The transition energy $2E_A$ reaches a minimum at $\delta = \pi$ (black arrow). **b**, Simplified diagram of the experimental set-up. A voltage-biased Josephson junction (orange box crossed diagonally, critical current 48 nA) is used as a spectrometer: it acts both as a microwave source and a detector. It is biased with a voltage source V_b in series with a resistor R_b . The a.c. Josephson current I_J (at frequency $\nu_J = 2eV_J/h$ set by the voltage V_J across the junction) is coupled through capacitor Σ (top) to a SQUID formed by an atomic point contact (magenta triangles) and an ancillary Josephson junction (black box crossed diagonally, critical current 1 μA , 20 times larger than the typical critical current of a one-atom aluminium contact). Magnetic flux ϕ threading the loop imposes a phase $\delta \approx \varphi = 2\pi\phi/\phi_0$ across the contact and determines the Andreev transition frequency of **a**. **c**, When biased at voltage V_J , the electrochemical potentials of the two electrodes of the spectrometer (black lines) are shifted by eV_J . The absorption of a photon at frequency $\nu_J = 2eV_J/h$ (orange wavy line) is accompanied by the transfer of a Cooper pair through the spectrometer (orange arrow labelled $2e$). **d**, Micrograph of the sample (at an angle of 45°) with spectrometer, suspended bridge to obtain the atomic contacts and SQUID Josephson junction. Components from **b** are shown arrowed.

¹Quantronics Group, Service de Physique de l'État Condensé (CNRS, URA 2464), IRAMIS, CEA-Saclay, 91191 Gif-sur-Yvette, France.

*These authors contributed equally to this work.

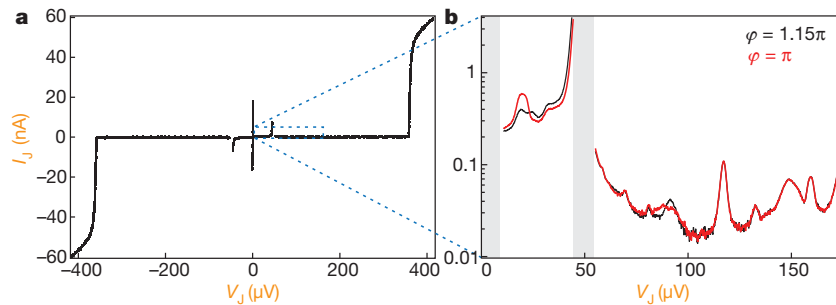


Figure 2 | $I_J(V_J)$ characteristics of the spectrometer coupled to the SQUID with atomic contact AC2. **a**, Data taken over a large range of V_J . The boxed region is shown magnified in **b**. **b**, The sub-gap current for two values of the reduced flux, ϕ . The grey regions at $V_J \leq 9 \mu\text{V}$ and on the right-hand side of the

peak at $V_J \approx 50 \mu\text{V}$ are not accessible because the biasing is unstable (see main text). The parts of the spectra that change with ϕ reveal energy absorption by the SQUID.

contacts, a simple system that allows direct quantitative comparison with theoretical predictions.

The principle of our experiment is described in Fig. 1b. An atomic contact obtained using a microfabricated, mechanically controllable break junction²¹ is placed in parallel with a tunnel Josephson junction to form a SQUID (superconducting quantum interference device). A second tunnel junction, the ‘spectrometer’, is used as an on-chip broadband microwave source and detector^{22–24}. It is coupled to the SQUID through an on-chip capacitor ($\sim 30 \text{ pF}$). The superconducting material for the junctions and atomic contact is aluminium ($\Delta \approx 180 \mu\text{eV}$; see Methods for fabrication details). A micrograph of the sample is shown in Fig. 1d. Both the spectrometer and the SQUID can be voltage-biased separately through on-chip inductor/capacitor (LC) filters (Supplementary Figs 1 and 2). The transmissions of the conduction channels of the atomic contact are determined by fitting the current-voltage characteristic of the SQUID with the theory of multiple Andreev reflections²⁵ (Supplementary Fig. 3). The SQUID geometry also allows phase biasing the atomic contact by applying a magnetic flux ϕ through the loop. Because the sum of the Josephson inductance of the SQUID tunnel junction ($\sim 310 \text{ pH}$) and the inductance of the SQUID loop ($\sim 20 \text{ pH}$) is much smaller than the typical atomic contact inductance ($\sim 3 \text{ nH}$), the phase difference across the atomic contact is $\delta \approx \phi = 2\pi\phi/\phi_0$.

When biased at a voltage V_J , the spectrometer undergoes Josephson oscillations and acts as a microwave current source at frequency $\nu_J = 2eV_J/h$. Microwave photons emitted by the spectrometer are

absorbed by the environment, which subsequently relaxes. The dissipated power P requires a d.c. current I_J to be supplied by the biasing circuit to satisfy power conservation, $P = I_J V_J$. Microscopically, this d.c. current is a result of inelastic Cooper-pair tunnelling: each time a photon is absorbed, a Cooper pair tunnels across the spectrometer insulating barrier^{26,27}, as in Fig. 1c. In the current-voltage $I_J(V_J)$ characteristic of the spectrometer junction, a transition of energy E is revealed as a d.c. current peak at $2eV_J = E$ with height $I_J = 2e\Gamma(E)$, where $\Gamma(E)$ is the photon absorption rate. Classically, this rate is related to the real part of the impedance seen by the spectrometer. The on-chip coupling capacitor and LC filtering are designed to keep the absorption rate due to the external environment low. Transitions such as the Andreev excitation ($|-\rangle \rightarrow |+\rangle$) at the energy $2eV_J = 2E_A(\delta, \tau)$ can be distinguished by their dependence on both the flux and the contact configuration.

Figure 2 shows the current-voltage characteristic $I_J(V_J)$ of the spectrometer for atomic contact AC2 (see below) at two values of the reduced flux ϕ . Several current peaks are visible below the voltage $\Delta/e \approx 180 \mu\text{V}$, which corresponds to the maximum excitation energy of interest for the Andreev transition, $2eV_J = 2\Delta$. Parts of the spectra change by as much as 200 pA as a function of reduced flux ϕ , revealing excitation of modes associated with the SQUID. Specifically, when going from $\phi = 1.15\pi$ (black line in Fig. 2b) to $\phi = \pi$ (red line), a prominent peak develops at a voltage bias $V_J = 20 \mu\text{V}$ and a peak at $\sim 90 \mu\text{V}$ broadens. Peaks which do not depend on the flux bias or the contact configuration, for example around $V_J = 150 \mu\text{V}$, are interpreted as resonances in the external electromagnetic environment and form a

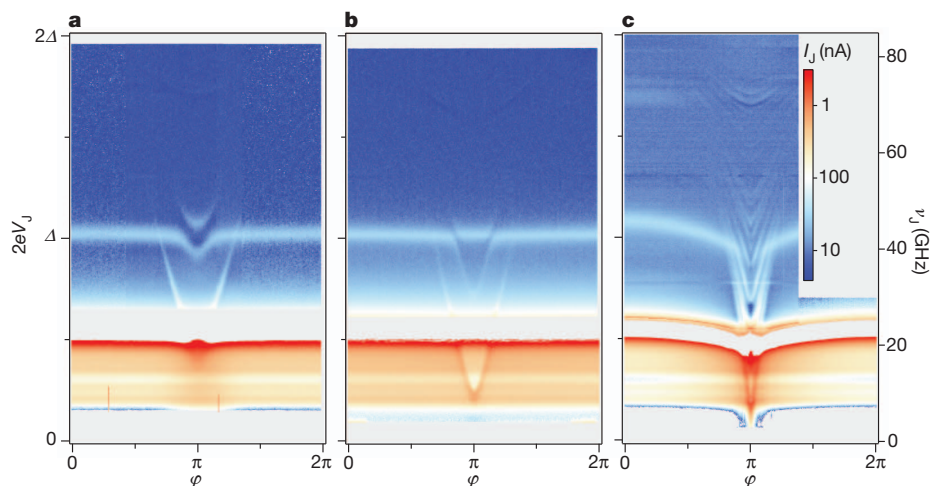


Figure 3 | Absorption spectra for three atomic contacts. **a**, AC1 (with two channels of transmission probability, 0.942 and 0.26); **b**, AC2 (with two channels of transmission probability, 0.985 and 0.37); **c**, AC3 (more than 20 channels). The colour encodes the current I_J through the spectrometer (colour

scale in **c** applies to all three panels), as a function of the reduced flux ϕ and of the bias voltage V_J . The right axis gives the spectrometer frequency ν_J associated with V_J .

background which is subtracted from the current–voltage curves (Supplementary Information section 2.2). In Fig. 2b there are no data in the two grey regions ($V_J \approx 50 \mu\text{V}$ and $V_J \leq 9 \mu\text{V}$) because the spectrometer voltage biasing is not stable (see Methods). The measured current decreases about an order of magnitude as the bias voltage is increased and passes through the zone of instability at $V_J \approx 50 \mu\text{V}$.

Spectra measured for the three different atomic contacts AC1, AC2 and AC3 are shown in Fig. 3a–c. In each spectrum, the current I_J through the spectrometer junction is plotted with the common colour scale of Fig. 3c. The vertical axes give the energy of photons emitted by the spectrometer in units of the bias voltage, $h\nu_J = 2eV_J$. The corresponding frequencies range from 0 to 85 GHz. The horizontal axes give the applied reduced flux, $\varphi \approx \delta$. There are no data in the grey regions where biasing is unstable. The contrast becomes fainter as the energy increases, except for a narrow band around 1.8Δ . The most remarkable features are the V-shaped transitions which fan out from $\varphi = \pi$ towards higher energies. AC3, which is a many-atom contact with about 20 conduction channels (Fig. 3c), has a multitude of well resolved V-shaped transitions. These transitions, which depend sensitively on the channel transmissions τ_i as well as φ , are the Andreev transitions. To confirm this, we plot with red lines in Fig. 4b (AC1) and Fig. 4d (AC2) the expected positions $2E_{A1}$ and $2E_{A2}$ of the Andreev transitions using equation (1) for the two highest transmission channels in each contact: AC1 (transmissions 0.942, 0.26) and AC2 (transmissions 0.985, 0.37) (see Supplementary Fig. 3). The lines match the observed transitions.

In addition to the Andreev transitions, there is in Fig. 3a–c a much brighter spectroscopic line ($I_J > 1 \text{ nA}$, colour scale red) common to all contacts which is located at 0.51Δ and hardly varies with flux for AC1 and AC2 but dips to 0.4Δ at $\varphi \approx \pi$ for AC3. It corresponds to the large peak at $V_J = 45 \mu\text{V}$ in Fig. 2, whose upper half falls in the region of instability. We identify it as the excitation of the plasma mode of the SQUID. This oscillator mode, formed by the SQUID Josephson inductance $L_S(\varphi)$ and its parallel capacitance C_S , resonates at frequency $\nu_p = (2\pi\sqrt{L_S C_S})^{-1} \approx 22 \text{ GHz}$ ($0.51\Delta/h$). The capacitance $C_S \approx 280 \text{ fF}$ is the sum of the SQUID and spectrometer capacitances. $L_S(\varphi)$ results from the parallel combination of three inductive elements: the atomic contact, the SQUID Josephson junction, and an on-chip inductor on the biasing line (Supplementary Fig. 1). The flux dependence of $L_S(\varphi)$ is negligible for the asymmetric SQUIDs (cases AC1, AC2) but results in a 0.1Δ amplitude modulation (4 GHz) of the plasma frequency for the large atomic-contact SQUID (case AC3). The energy $h\nu_p$ associated with the plasma frequency ν_p for AC1 and AC2 is plotted in Fig. 4b and d, respectively, as blue lines, and agrees with the experimental data. The abrupt decrease in spectrometer signal above the plasma frequency (Fig. 2b) is due to the shunting of emitted microwaves by the capacitance C_S .

The combination of the Andreev and plasma degrees of freedom leads to a double ladder energy diagram, as shown in Fig. 4a. The states are labelled by $|\sigma, n\rangle$, where $\sigma = \pm$ accounts for the Andreev pair state and n is the plasmon number. The data are well explained by considering transitions only from the initial state $|- , 0\rangle$. The Andreev transition ($|- , 0\rangle \rightarrow |+, 0\rangle$) at $2eV_J = 2E_A$ is indicated by the red arrow and the plasma transition ($|- , 0\rangle \rightarrow |-, 1\rangle$) at $2eV_J = h\nu_p$ by the blue arrow.

In the spectrum of each contact, there is another resonance near 1.02Δ , which is similar in shape to the plasma transition but at twice the energy and of smaller amplitude ($\sim 100 \text{ pA}$). This corresponds to the second harmonic of the plasma transition, $2eV_J = 2h\nu_p$ (Figs 4b and d, blue dashed line), in which each Cooper pair tunnelling through the spectrometer emits two photons of energy $h\nu_p$. This two-photon plasma transition $|- , 0\rangle \rightarrow |-, 2\rangle$ is represented by the blue dashed arrow in the energy ladder, Fig. 4a. It is also possible to simultaneously excite the Andreev transition and the plasma mode (Fig. 4a, purple dashed arrow). This type of transition $|- , 0\rangle \rightarrow |+, 1\rangle$, at $2eV_J = 2E_A + h\nu_p$, is observed in the spectra, Fig. 4b and d, as a replica of the Andreev transition, shifted up by the plasma energy (purple dashed line). These transitions agree with the data everywhere except where two

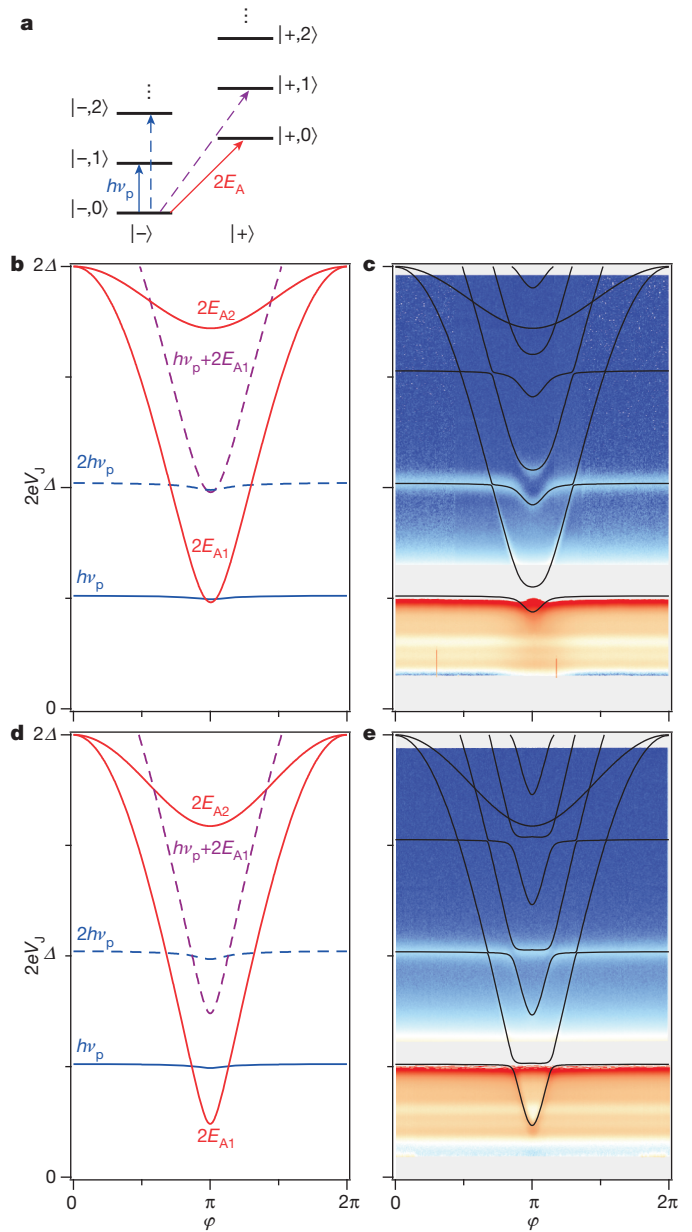


Figure 4 | Interpretation of the absorption spectra. **a**, Energy spectrum diagrams for a single channel: each state is labelled $|\sigma, n\rangle$ for the Andreev pair in the ground ($-$) or excited ($+$) state and n photons in the plasma mode. **b–e**, Predicted transitions between the states shown in **a** using the bare Andreev and plasma energies (**b, d**) or by diagonalization of the full Hamiltonian (**c, e**), for contacts AC1 (**b, c**) and AC2 (**d, e**). **b, d**, Red lines, transition energies $2E_A$ predicted from the channel transmissions (equation (1)): $2E_{A1}$ for first channel and $2E_{A2}$ for second channel). Blue lines, excitation energy for the plasma mode $h\nu_p$. Blue dashed line, two-photon plasma mode process $2h\nu_p$. Purple dashed line, two-photon process $2E_A + h\nu_p$ exciting both the Andreev and plasma transitions. **c, e**, Black lines, calculated transition lines superimposed on the data.

such two-photon processes coincide, near $\varphi = \pi$ and $2eV_J \approx 1.02\Delta$. There one observes a level repulsion (Fig. 3a) or an avoided crossing (Fig. 3b), depending on the relative position of the undressed states. In the spectra the region of instability obscures the hybridization effects at energy $2eV_J = h\nu_p$ but line traces slightly below confirm their existence (see Supplementary Information).

The experimental spectra are well described by a model based on the Andreev Hamiltonian¹¹ (Supplementary Information section 2.3). The eigenenergies of the SQUID Hamiltonian are determined by perturbation analysis and numerical diagonalization. The resulting transition

energies are shown as black lines in Fig. 4c and e. Only crossings of transition lines involving the same number of photons show significant hybridization, in good agreement with the data. The rich structure predicted in the top part of the spectrum is not visible in the experiment because of the shunting by the SQUID capacitor. A quantitative description of the intensity and width of the transitions would require taking into account the coupling to the detector and the sources of dissipation.

Our results show that in addition to the phase difference, each conduction channel of a Josephson weak link possesses an internal degree of freedom similar to a spin-1/2. This Andreev pseudo-spin is unique as a microscopic degree of freedom intrinsically coupled to a superconducting circuit and whose energy is tunable over a wide range. Theoretical proposals for an Andreev qubit are based either directly on this pseudo-spin¹¹ or on the actual spin of quasiparticles trapped in the Andreev levels^{12,13,16}. Their implementation requires reducing external sources of decoherence, something that could be achieved, in the circuit quantum electrodynamics approach, by integrating a superconducting atomic contact in a high-quality resonator^{28,29}. Finally, in hybrid systems where spin-orbit and Zeeman interactions are also present, Andreev levels give rise to Majorana states whose detection is currently the subject of intense study³⁰.

METHODS SUMMARY

The sample is mounted in a bending mechanism (Supplementary Fig. 2d) anchored to the mixing chamber of a dilution refrigerator at 30 mK and housed inside a superconducting shield to reduce magnetic interference. Two microwave launchers connect it to the biasing and measuring lines which are heavily filtered. An electrically shielded small superconducting coil located directly above the sample is used to apply magnetic flux. A pusher actuated by a room temperature d.c. motor bends the sample and modifies the atomic contact configuration. The atomic contacts, tunnel junctions, and on-chip filters (alumina dielectric) are fabricated by electron-beam lithography and evaporation. Tunnel junctions are formed by double-angle evaporation and oxidation and have a bare plasma frequency of 14 GHz. Measurements of the SQUID and spectrometer current-voltage characteristics are made at low frequency (10–100 Hz) with room-temperature amplification. When the differential conductance of the spectrometer is smaller than $-1/R_b$, with R_b the biasing resistor of the spectrometer (see Fig. 1b), such as on the negative-slope side of the first plasma peak, biasing is unstable. This results in the absence of data in the grey regions above the plasma transition in Figs 2 and 3. At low voltages, there is another instability due to retrapping to the zero-voltage state. The peaks in the current-voltage curves which do not depend on the flux are subtracted from the measured spectra in the region $V_1 > 50 \mu\text{V}$ as described in Supplementary Information section 2.2. The theoretical spectra of Fig. 4c and e are obtained by numerical diagonalization of the Hamiltonian describing both the Andreev states and the plasma mode, which are coupled because they share the phase across the SQUID Josephson junction (Supplementary Information section 2.3).

Received 5 March; accepted 16 May 2013.

- Josephson, B. D. Possible new effects in superconductive tunnelling. *Phys. Lett.* **1**, 251–253 (1962).
- Busch, S. *et al.* Measurements of T_1 -relaxation in *ex vivo* prostate tissue at 132 μT . *Magn. Reson. Med.* **67**, 1138–1145 (2012).
- Erik Lucero *et al.* High-fidelity gates in a single Josephson qubit. *Nature Phys.* **8**, 719–723 (2012).
- Wendin, G. & Shumeiko, V. S. Quantum bits with Josephson junctions. *Low Temp. Phys.* **33**, 724–744 (2007).
- Fink, J. M. *et al.* Climbing the Jaynes–Cummings ladder and observing its nonlinearity in a cavity QED system. *Nature* **454**, 315–318 (2008).
- Hofheinz, M. *et al.* Synthesizing arbitrary quantum states in a superconducting resonator. *Nature* **459**, 546–549 (2009).
- Kulik, I. O. Macroscopic quantization and proximity effect in S–N–S junctions. *Sov. Phys. JETP* **30**, 944–950 (1970).

- Furusaki, A. & Tsukada, M. Dc Josephson effect and Andreev reflection. *Solid State Commun.* **78**, 299–302 (1991).
- Beenakker, C. W. J. & van Houten, H. Josephson current through a superconducting quantum point contact shorter than the coherence length. *Phys. Rev. Lett.* **66**, 3056–3059 (1991).
- Bagwell, P. F. Suppression of the Josephson current through a narrow, mesoscopic, semiconductor channel by a single impurity. *Phys. Rev. B* **46**, 12573–12586 (1992).
- Zazunov, A., Shumeiko, V. S., Bratus', E. N., Lantz, J. & Wendin, G. Andreev level qubit. *Phys. Rev. Lett.* **90**, 087003 (2003).
- Chitchev, N. M. & Nazarov, Yu. V. Andreev quantum dots for spin manipulation. *Phys. Rev. Lett.* **90**, 226806 (2003).
- Padurariu, C. & Nazarov, Yu. V. Spin blockade qubit in a superconducting junction. *Europhys. Lett.* **100**, 57006–57011 (2012).
- Bardeen, J., Cooper, L. N. & Schrieffer, J. R. Theory of superconductivity. *Phys. Rev.* **108**, 1175–1204 (1957).
- Della Rocca, M. L. *et al.* Measurement of the current-phase relation of superconducting atomic contacts. *Phys. Rev. Lett.* **99**, 127005 (2007).
- Zgirski, M. *et al.* Evidence for long-lived quasiparticles trapped in superconducting point contacts. *Phys. Rev. Lett.* **106**, 257003 (2011).
- Deacon, R. S. *et al.* Tunneling spectroscopy of Andreev energy levels in a quantum dot coupled to a superconductor. *Phys. Rev. Lett.* **104**, 076805 (2010).
- Pillet, J.-D. *et al.* Andreev bound states in supercurrent-carrying carbon nanotubes revealed. *Nature Phys.* **6**, 965–969 (2010).
- Morpurgo, A. F., Baselmans, J. J. A., van Wees, B. J. & Klapwijk, T. M. Energy spectroscopy of the Josephson supercurrent. *J. Low Temp. Phys.* **118**, 637–651 (2000).
- Fuechsle, M. *et al.* Effect of microwaves on the current-phase relation of superconductor normal-metal superconductor Josephson junctions. *Phys. Rev. Lett.* **102**, 127001 (2009).
- van Ruitenbeek, J. M. *et al.* Adjustable nanofabricated atomic size contacts. *Rev. Sci. Instrum.* **67**, 108–111 (1996).
- Edstam, J. & Olsson, H. K. Josephson broadband spectroscopy to 1 THz. *Appl. Phys. Lett.* **64**, 2733–2735 (1994).
- Leppäkangas, J., Thuneberg, E., Lindell, R. & Hakonen, P. Tunneling of Cooper pairs across voltage-biased asymmetric single-Cooper-pair transistors. *Phys. Rev. B* **74**, 054504 (2006).
- Billangeon, P.-M., Pierre, F., Bouchiat, H. & Deblock, R. Very high frequency spectroscopy and tuning of a single-Cooper-pair transistor with an on-chip generator. *Phys. Rev. Lett.* **98**, 126802 (2007).
- Scheer, E., Joyez, P., Esteve, D., Urbina, C. & Devoret, M. H. Conduction channel transmissions of atomic-size aluminum contacts. *Phys. Rev. Lett.* **78**, 3535–3538 (1997).
- Holst, T., Esteve, D., Urbina, C. & Devoret, M. H. Effect of a transmission line resonator on a small capacitance tunnel junction. *Phys. Rev. Lett.* **73**, 3455–3458 (1994).
- Hofheinz, M. *et al.* Bright side of the Coulomb blockade. *Phys. Rev. Lett.* **106**, 217005 (2011).
- Romero, G., Lizuain, I., Shumeiko, V. S., Solano, E. & Bergeret, F. S. Circuit quantum electrodynamics with a superconducting quantum point contact. *Phys. Rev. B* **85**, 180506 (2012).
- Sköldberg, J., Löfwander, T., Shumeiko, V. S. & Fogelström, M. Spectrum of Andreev bound states in a molecule embedded inside a microwave-excited superconducting junction. *Phys. Rev. Lett.* **101**, 087002 (2008).
- Mourik, V. *et al.* Signatures of Majorana fermions in hybrid superconductor-semiconductor nanowire devices. *Science* **336**, 1003–1007 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge technical assistance from P. Sénat and P.-F. Orfila, theoretical input from M. Houzet, help in the experiments from L. Tosi, and discussions with V. Shumeiko, A. Levy-Yeyati and within the Quantronics group. This work was supported by ANR contracts DOCFLUC and MASH, and by C'Nano. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007–2013) under REA grant agreement no. PIFI-GA-2011-298415.

Author Contributions All authors designed the experiment, L.B. and Ç.Ö.G. fabricated the sample, L.B., Ç.Ö.G., H.P. and C.U. carried out the measurements and analysed the data, and all authors contributed to the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.U. (cristian.urbina@cea.fr).

Sequential deposition as a route to high-performance perovskite-sensitized solar cells

Julian Burschka^{1*}, Norman Pellet^{1,2*}, Soo-Jin Moon¹, Robin Humphry-Baker¹, Peng Gao¹, Mohammad K. Nazeeruddin¹ & Michael Grätzel¹

Following pioneering work¹, solution-processable organic–inorganic hybrid perovskites—such as $\text{CH}_3\text{NH}_3\text{PbX}_3$ ($\text{X} = \text{Cl}, \text{Br}, \text{I}$)—have attracted attention as light-harvesting materials for mesoscopic solar cells^{2–15}. So far, the perovskite pigment has been deposited in a single step onto mesoporous metal oxide films using a mixture of PbX_2 and $\text{CH}_3\text{NH}_3\text{X}$ in a common solvent. However, the uncontrolled precipitation of the perovskite produces large morphological variations, resulting in a wide spread of photovoltaic performance in the resulting devices, which hampers the prospects for practical applications. Here we describe a sequential deposition method for the formation of the perovskite pigment within the porous metal oxide film. PbI_2 is first introduced from solution into a nanoporous titanium dioxide film and subsequently transformed into the perovskite by exposing it to a solution of $\text{CH}_3\text{NH}_3\text{I}$. We find that the conversion occurs within the nanoporous host as soon as the two components come into contact, permitting much better control over the perovskite morphology than is possible with the previously employed route. Using this technique for the fabrication of solid-state mesoscopic solar cells greatly increases the reproducibility of their performance and allows us to achieve a power conversion efficiency of approximately 15 per cent (measured under standard AM1.5G test conditions on solar zenith angle, solar light intensity and cell temperature). This two-step method should provide new opportunities for the fabrication of solution-processed photovoltaic cells with unprecedented power

conversion efficiencies and high stability equal to or even greater than those of today's best thin-film photovoltaic devices.

We prepared mesoporous TiO_2 (anatase) films by spin-coating a solution of colloidal anatase particles onto a 30-nm-thick compact TiO_2 underlayer. The underlayer was deposited by aerosol spray pyrolysis on a transparent-conducting-oxide-coated glass substrate acting as the electric front contact of the solar cell. Lead iodide (PbI_2) was then introduced into the TiO_2 nanopores by spin-coating a 462 mg ml^{-1} ($\sim 1 \text{ M}$) solution of PbI_2 in *N,N*-dimethylformamide (DMF) kept at 70°C . The use of such a high PbI_2 concentration is critical to obtaining the high loading of the mesoporous TiO_2 films required to fabricate solar cells of the highest performance. Further experimental details are provided in Methods.

Figure 1a presents a cross-sectional scanning electron microscopy (SEM) image of the thus-prepared film. The absence of any PbI_2 crystals protruding from the surface of the mesoporous anatase layer shows that our infiltration method leads to a structure in which the PbI_2 is entirely contained within the nanopores of the TiO_2 film.

Dipping the $\text{TiO}_2/\text{PbI}_2$ composite film into a solution of $\text{CH}_3\text{NH}_3\text{I}$ in 2-propanol (10 mg ml^{-1}) changes its colour immediately from yellow to dark brown, indicating the formation of $\text{CH}_3\text{NH}_3\text{PbI}_3$. We monitored the dynamics of the formation of the perovskite by optical absorption, emission and X-ray diffraction (XRD) spectroscopy. Figure 1b shows that the increase over time of the perovskite absorption at 550 nm is

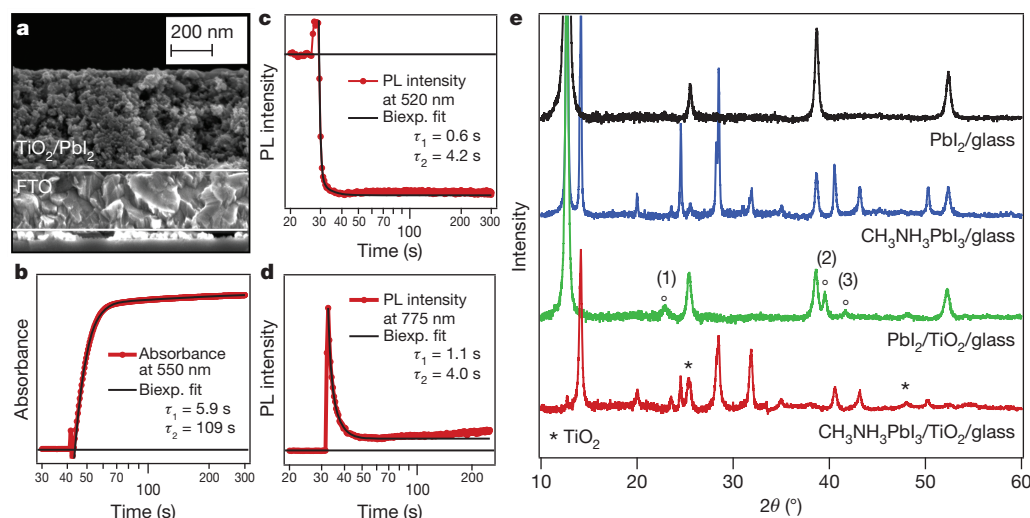


Figure 1 | Transformation of PbI_2 into $\text{CH}_3\text{NH}_3\text{PbI}_3$ within the nanopores of a mesoscopic TiO_2 film. **a**, Cross-sectional SEM of a mesoporous TiO_2 film infiltrated with PbI_2 . FTO, fluorine-doped tin oxide. **b**, Change in absorbance at 550 nm of such a film monitored during the transformation. **c**, Change in photoluminescence (PL) intensity at 520 nm monitored during the

transformation. Excitation at 460 nm. **d**, Change in photoluminescence intensity at 775 nm monitored during the transformation. Excitation at 660 nm. **e**, X-ray diffraction spectra of PbI_2 on glass and porous $\text{TiO}_2/\text{glass}$ before and after the transformation. The dipping time was 60 s in both cases. The plot shows the X-ray intensity as a function of 2θ (twice the diffraction angle).

¹Laboratory of Photonics and Interfaces, Department of Chemistry and Chemical Engineering, Swiss Federal Institute of Technology, Station 6, CH-1015 Lausanne, Switzerland. ²Max-Planck-Institute for Solid-State Research, Heisenbergstraße 1, D-70569 Stuttgart, Germany.

*These authors contributed equally to this work.

practically complete within a few seconds of exposing the PbI_2 -loaded TiO_2 film to the $\text{CH}_3\text{NH}_3\text{I}$ solution. A small additional increase in the absorbance, occurring on a timescale of 100 s and contributing only a few per cent to the total increase of the signal, is attributed to morphological changes producing enhanced light scattering. The conversion is accompanied by a quenching of the PbI_2 emission at 425 nm (Fig. 1c) and a concomitant increase in the perovskite luminescence at 775 nm (Fig. 1d). The latter emission passes through a maximum before decreasing to a stationary value. This decrease results from self-absorption of the luminescence by the perovskite formed during the reaction with $\text{CH}_3\text{NH}_3\text{I}$. The traces were fitted to a biexponential function yielding the decay times stated in Fig. 1b–d. We note that the increase in the emission intensity before the quenching in Fig. 1c is an optical artefact that results from opening the sample compartment to add the $\text{CH}_3\text{NH}_3\text{I}$ solution.

The green and red curves in Fig. 1e show X-ray powder diffraction spectra measured before and, respectively, after the $\text{TiO}_2/\text{PbI}_2$ nanocomposite film is brought into contact with the $\text{CH}_3\text{NH}_3\text{I}$ solution. For comparison, we spin-coated the PbI_2 also on a flat glass substrate and exposed the resulting film to a $\text{CH}_3\text{NH}_3\text{I}$ solution in the same manner as the $\text{TiO}_2/\text{PbI}_2$ nanocomposite. On the basis of literature data, the PbI_2 deposited by spin-coating from DMF solution crystallizes in the form of the hexagonal 2H polytype, the most common PbI_2 modification (Inorganic Crystal Structure Database, collection code 68819; <http://www.fiz-karlsruhe.com/icsd.html>). Moreover, the results show that on a flat glass substrate, crystals grow in a preferential orientation along the *c* axis, hence the appearance of only four diffraction peaks, corresponding to the (001), (002), (003) and (004) lattice planes (Fig. 1e, black curve). For the PbI_2 loaded on a mesoporous TiO_2 film (Fig. 1e, green curve), we find three additional diffraction peaks that do not originate from TiO_2 , suggesting that the anatase scaffold induces a different orientation for the PbI_2 crystal growth. The peaks labelled (2) and (3) in Fig. 1e can be attributed to the (110) and (111) lattice planes of the 2H polytype. Peak (1) is assigned to a different PbI_2 variant, whose identification is beyond the scope of this report in view of the large number of polytypes that have been reported for PbI_2 (ref. 16).

During the reaction with $\text{CH}_3\text{NH}_3\text{I}$, we observe the appearance of a series of new diffraction peaks that are in good agreement with literature data on the tetragonal phase of the $\text{CH}_3\text{NH}_3\text{PbI}_3$ perovskite¹⁷. However, when PbI_2 is deposited on a flat film (Fig. 1e, blue curve) the conversion to perovskite on exposure to the $\text{CH}_3\text{NH}_3\text{I}$ solution is incomplete; a large amount of unreacted PbI_2 remained even after a dipping time of 45 min. This agrees with the observation that the $\text{CH}_3\text{NH}_3\text{I}$ insertion hardly proceeds beyond the surface of thin PbI_2 films, and that the complete transformation of the crystal structure requires several hours¹⁸. A caveat associated with such long conversion times is that the perovskite dissolves in the methylammonium iodide solution over longer periods, hampering the transformation.

In striking contrast to the behaviour of thin films of PbI_2 deposited on a flat support, the conversion of PbI_2 nanocrystals in the mesoporous TiO_2 film is practically complete on a timescale of seconds, as is evident from the immediate disappearance of its most intense diffraction peak (the (001) peak) and the concomitant appearance of the XRD reflections for the tetragonal perovskite. When the PbI_2 crystals are contained within the mesoporous TiO_2 scaffold, their size is limited to ~ 22 nm by the pore size of the host. Notably, we find that confining the PbI_2 crystals to such a small size drastically enhances their rate of conversion to perovskite, which is complete within a few seconds of their coming into contact with the methylammonium iodide solution. However, when deposited on a flat surface, larger PbI_2 crystallites in the size range of 50–200 nm are formed, resulting in incomplete conversion of the PbI_2 on exposure to $\text{CH}_3\text{NH}_3\text{I}$, as shown by XRD. The SEM images of such a film that are depicted in Supplementary Fig. 1e, f show, however, that the perovskite produced by the sequential deposition technique adopts a morphology similar to that of the PbI_2 precursor. Supplementary Fig. 1 also shows that large crystals of $\text{CH}_3\text{NH}_3\text{PbI}_3$ with a

wide range of sizes are formed when the perovskite is deposited in a single step from a solution of $\text{CH}_3\text{NH}_3\text{I}$ and PbI_2 in γ -butyrolactone or DMF.

A key finding of the present work is that the confinement of the PbI_2 within the nanoporous network of the TiO_2 film greatly facilitates its conversion to the perovskite pigment. Moreover, the mesoporous scaffold of the host forces the perovskite to adopt a confined nanomorphology.

The literature contains several examples in which a two-step procedure is used to fabricate nanostructures that are not easily, or not at all, accessible by a direct synthetic route. Ion exchange reactions have, for example, been used to convert dispersed II–V semiconductor nanocrystals into the corresponding III–V analogues while preserving particle size and distribution as well as the initial nanomorphology^{19–21}. As reported, the thermodynamic driving force of such a reaction is the difference in bulk lattice energy for the two materials, and the initial crystal lattice serves as a template for the formation of the desired compound. As for PbI_2 , the insertion of the organic cation is facilitated through the layered PbI_2 structure, which consists of three spatially repeating planes, I–Pb–I (ref. 16). Numerous literature reports show that strong intralayer chemical bonding, as well as only weak interlayer van der Waals interactions, allows the easy insertion of guest molecules between these layers^{22–24}. In our case, the large energy of formation of the hybrid perovskite, combined with the nanoscopic morphology of the PbI_2 precursor, which greatly enhances the reaction kinetics, finally enables the transformation to be completed within seconds.

We used the sequential deposition technique to fabricate mesoscopic solar cells employing the triarylamine derivative 2,2',7,7'-tetrakis(*N*, *N*-di-*p*-methoxyphenylamine)-9,9'-spirobifluorene (spiro-MeOTAD) (Supplementary Fig. 2) as a hole-transporting material (HTM). We note that, following a recently reported concept²⁵, we use a Co(III) complex as a p-type dopant for the HTM at a molar doping level of 10% to ensure a sufficient conductivity and low series resistance. Figure 2 shows a cross-sectional SEM picture of a typical device. The mesoporous TiO_2 film had an optimized thickness of around 350 nm and was infiltrated with the perovskite nanocrystals using the above-mentioned two-step procedure. The HTM was subsequently deposited by spin coating. It penetrates into the remaining available pore volume and forms a 100-nm-thick layer on top of the composite structure. A thin gold layer was thermally evaporated under vacuum onto the HTM, forming the back contact of the device.

We measured the current density (*J*)–voltage (*V*) characteristics of the solar cells under simulated air mass 1.5 global (AM1.5G) solar irradiation and in the dark. Figure 3a shows *J*–*V* curves measured at a light intensity of 95.6 mW cm^{-2} for a typical device. From this, we derive values for the short-circuit photocurrent (J_{sc}), the open-circuit voltage (V_{oc}) and the fill factor of, respectively, 17.1 mA cm^{-2} , 992 mV and 0.73, yielding a solar-to-electric power conversion efficiency (PCE) of 12.9% (Table 1). Statistical data on a larger batch of ten photovoltaic devices is shown in Supplementary Table 1. From the

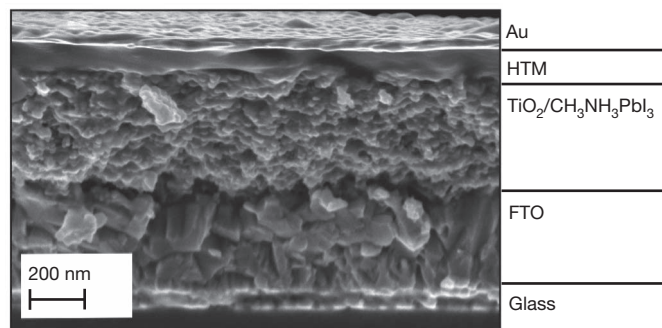


Figure 2 | Cross-sectional SEM of a complete photovoltaic device. Note that the thin TiO_2 compact layer present between the FTO and the mesoscopic composite is not resolved in the SEM image.

average PCE value of $12.0\% \pm 0.5\%$ and the small standard deviation, we infer that photovoltaics with excellent performance and high reproducibility can be realized using the method reported here.

Figure 3b shows the incident-photon-to-current conversion efficiency (IPCE), or external quantum efficiency, spectrum for the perovskite cell. Generation of photocurrent starts at 800 nm, in agreement with the bandgap of the $\text{CH}_3\text{NH}_3\text{PbI}_3$, and reaches peak values of over 90% in the short-wavelength region of the visible spectrum. Integrating the overlap of the IPCE spectrum with the AM1.5G solar photon flux yields a current density of 18.4 mA cm^{-2} , which is in excellent agreement with the measured photocurrent density, extrapolated to 17.9 mA cm^{-2} at the standard solar AM1.5G intensity of 100 mW cm^{-2} . This confirms that any mismatch between the simulated sunlight and the AM1.5G standard is negligibly small. Comparison with the absorbance or light-harvesting efficiency (LHE) depicted in Fig. 3c reveals that the low IPCE values in the range of 600–800 nm result from the smaller absorption of the perovskite in this spectral region. This is also reflected in the spectrum of the internal quantum efficiency, or absorbed-photon-to-current conversion efficiency (APCE), which can be derived from the IPCE and LHE and is shown in Fig. 3d. The APCE is greater than 90% over the whole visible region, without correction for reflective losses, indicating that the device achieves near-unity quantum yield for the generation and collection of charge carriers.

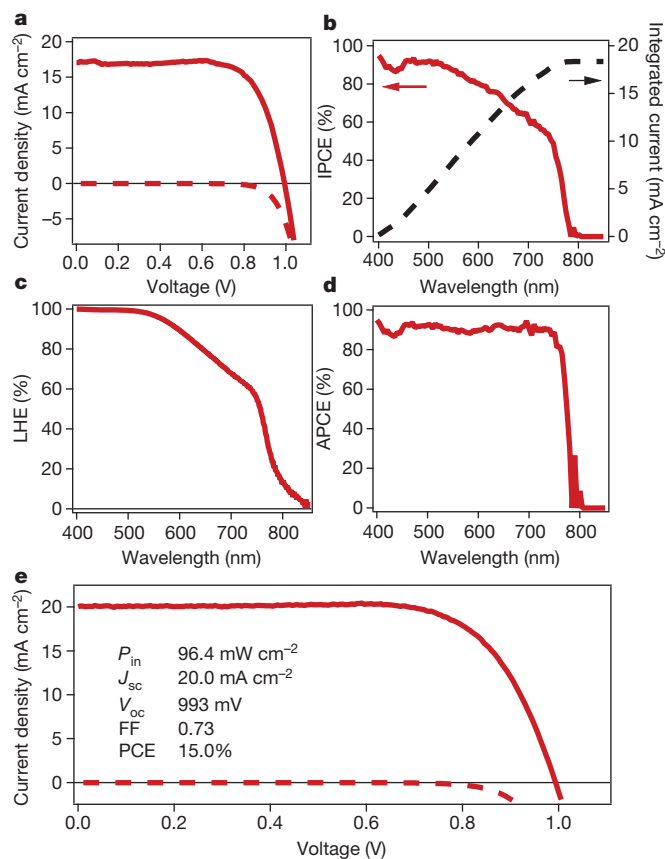


Figure 3 | Photovoltaic device characterization. **a**, J - V curves for a photovoltaic device measured at a simulated AM1.5G solar irradiation of 95.6 mW cm^{-2} (solid line) and in the dark (dashed line). **b**, IPCE spectrum. The right-hand axis indicates the integrated photocurrent that is expected to be generated under AM1.5G irradiation. **c**, LHE spectrum. **d**, APCE spectrum derived from the IPCE and LHE. **e**, J - V curves for a best-performing cell measured at a simulated AM1.5G solar irradiation of 96.4 mW cm^{-2} (solid line) and in the dark (dashed line). The device was fabricated using slightly modified deposition conditions (Methods). FF, fill factor.

Table 1 | Photovoltaics performance at different light intensities

Intensity (mW cm^{-2})	J_{sc} (mA cm^{-2})	V_{oc} (mV)	Fill factor	PCE (%)
9.3	1.7	901	0.77	12.6
49.8	8.9	973	0.75	13.0
95.6	17.1	992	0.73	12.9

In an attempt to increase the loading of the perovskite absorber on the TiO_2 structure and to obviate the lack of absorption in the long-wavelength region of the spectrum, we slightly modified the conditions for the deposition of the PbI_2 precursor as well as the transformation reaction. Details are provided in Methods. The J - V characteristics of the best-performing cell of the series that was fabricated in this manner are depicted in Fig. 3e. From this data, we derive values of 20.0 mA cm^{-2} , 993 mV and 0.73 for J_{sc} , V_{oc} and the fill factor, respectively, yielding a PCE of 15.0% measured at a light intensity of $P_{\text{in}} = 96.4 \text{ mW cm}^{-2}$. To the best of our knowledge, this is the highest power conversion efficiency reported so far for organic or hybrid inorganic-organic solar cells and one of the highest for any solution-processed photovoltaic device. Several solar cells with PCEs between 14% and 15% were fabricated. One of these devices was sent to an accredited photovoltaic calibration laboratory for certification, confirming a power conversion efficiency of 14.14% measured under standard AM1.5G reporting conditions. Detailed photovoltaics data for this device can be found in Supplementary Fig. 3. Compared with the devices from which we took the data shown in Fig. 3a and Supplementary Table 1, these top-performance devices benefit from a significantly higher photocurrent. We attribute this trend to the increased loading of the porous TiO_2 film with the perovskite pigment and to increased light scattering, improving the long-wavelength response of the cell. The increase in light scattering is likely to result from the additional pre-wetting step that was used for the top-performance devices. The pre-wetting locally decreases the methylammonium iodide concentration, inducing the growth of larger crystals. Detailed studies that aim to identify the key role of the different parameters during the sequential deposition are ongoing.

To test the stability of the perovskite-based photovoltaics prepared using the aforementioned procedure, we subjected a sealed cell to long-term light soaking at a light intensity of $\sim 100 \text{ mW cm}^{-2}$ and a temperature of 45°C . The device was encapsulated under argon and maintained at the optimal electric power output during the ageing using maximum-power-point tracking. We found a very promising long-term stability: the photovoltaic device maintained more than 80% of its initial PCE after a period of 500 h (Supplementary Fig. 4). Also, it is notable that we do not observe any change in the short-circuit photocurrent, indicating that there is no photodegradation of the perovskite light harvester. The decrease in PCE is therefore due only to a decrease in both the open-circuit potential and the fill factor, and the similar shape of the two decays suggests that they are linked to the same degradation mechanism. The change in these two parameters is mainly due to a decrease in the shunt resistance, as is apparent from Supplementary Fig. 5, where J - V curves of the device before and after the ageing process are shown.

The sequential deposition method for the fabrication of perovskite-sensitized mesoscopic solar cells introduced here provides a means to achieve excellent photovoltaic performance with high reproducibility. The power conversion efficiency of 15% achieved with the best device is amongst the highest for solution-processed photovoltaics and sets a new record for organic or hybrid inorganic-organic solar cells in general. Our findings open new routes for the fabrication of perovskite-based photovoltaic devices, because other preformed metal halide mesostructures may be converted into the desired perovskite by the simple insertion reaction detailed here. On the basis of our results, we believe that this new class of mesoscopic solar cells will find widespread application and will eventually lead to devices that rival conventional silicon-based photovoltaics.

METHODS SUMMARY

Device fabrication. Patterned transparent conducting oxide substrates were coated with a TiO₂ compact layer by aerosol spray pyrolysis. A 350-nm-thick mesoporous TiO₂ layer composed of 20-nm-sized particles was then deposited by spin coating. The mesoporous TiO₂ films were infiltrated with PbI₂ by spin-coating a PbI₂ solution in DMF (462 mg ml⁻¹) that was kept at 70 °C. After drying, the films were dipped in a solution of CH₃NH₃I in 2-propanol (10 mg ml⁻¹) for 20 s and rinsed with 2-propanol. After drying, the HTM was deposited by spin-coating a solution of spiro-MeOTAD, 4-tert-butylpyridine, lithium bis(trifluoromethylsulfonyl)imide and tris(2-(1H-pyrazol-1-yl)-4-tert-butylpyridine)cobalt(III) bis(trifluoromethylsulfonyl)imide in chlorobenzene. Gold (80 nm) was thermally evaporated on top of the device to form the back contact. For the fabrication of the best-performing devices, slightly modified conditions were used. First, PbI₂ was spin-cast for 5 s instead of 90 s and, second, the samples were subjected to a 'pre-wetting' by dipping in 2-propanol for 1–2 s before being dipped in the solution of CH₃NH₃I and 2-propanol. For all measurements, devices were equipped with a 0.285-cm² metal aperture to define the active area.

Long-term stability. Devices were sealed in argon and subjected to constant light soaking at approximately 100 mW cm⁻² using an array of light-emitting diodes. During the testing, the devices were maintained at their maximum power point and a temperature of about 45 °C. Current density/voltage curves were recorded automatically every 2 h.

Optical spectroscopy. Mesoporous TiO₂ films were deposited on microscope glass slides and infiltrated with PbI₂ following the above-mentioned procedure. The samples were then placed vertically in a cuvette with a path length of 10 mm. The solution of CH₃NH₃I was then rapidly injected into the cuvette while either the photoluminescence or the optical transmission was monitored.

Full Methods and any associated references are available in the online version of the paper.

Received 3 April; accepted 29 May 2013.

Published online 10 July 2013.

- Kojima, A., Teshima, K., Shirai, Y. & Miyasaka, T. Organometal halide perovskites as visible-light sensitizers for photovoltaic cells. *J. Am. Chem. Soc.* **131**, 6050–6051 (2009).
- Hagfeldt, A., Boschloo, G., Sun, L., Kloo, L. & Pettersson, H. Dye-sensitized solar cells. *Chem. Rev.* **110**, 6595–6663 (2010).
- Im, J.-H. *et al.* 6.5% efficient perovskite quantum-dot-sensitized solar cell. *Nanoscale* **3**, 4088–4093 (2011).
- Kim, H.-S. *et al.* Lead iodide perovskite sensitized all-solid-state submicron thin film mesoscopic solar cell with efficiency exceeding 9%. *Sci. Rep.* **2**, 591 (2012).
- Lee, M. M., Teuscher, J., Miyasaka, T., Murakami, T. N. & Snaith, H. J. Efficient hybrid solar cells based on meso-superstructured organometal halide perovskites. *Science* **338**, 643–647 (2012).
- Etgar, L. *et al.* Mesoscopic CH₃NH₃PbI₃/TiO₂ heterojunction solar cells. *J. Am. Chem. Soc.* **134**, 17396–17399 (2012).
- Im, J.-H., Chung, J., Kim, S.-J. & Park, N.-G. Synthesis, structure, and photovoltaic property of a nanocrystalline 2H perovskite-type novel sensitizer (CH₃CH₂NH₃)PbI₃. *Nanoscale Res. Lett.* **7**, 353 (2012).
- Edri, E., Kirmayer, S., Cahen, D. & Hodes, G. High open-circuit voltage solar cells based on organic-inorganic lead bromide perovskite. *Phys. Chem. Lett.* **4**, 897–902 (2013).
- Crossland, E. J. W. *et al.* Mesoporous TiO₂ single crystals delivering enhanced mobility and optoelectronic device performance. *Nature* **495**, 215–219 (2013).
- Noh, J. H., Im, S. H., Heo, J. H., Mandal, T. N. & Seok, S. I. Chemical management for colorful, efficient, and stable inorganic–organic hybrid nanostructured solar cells. *Nano Lett.* **13**, 1764–1769 (2013).
- Cai, B., Xing, Y., Yang, Z., Zhang, W.-H. & Qiu, J. High performance hybrid solar cells sensitized by organolead halide perovskites. *Energy Environ. Sci.* **6**, 1480–1485 (2013).
- Qui, J. *et al.* All-solid-state hybrid solar cells based on a new organometal halide perovskite sensitizer and one-dimensional TiO₂ nanowire arrays. *Nanoscale* **5**, 3245–3248 (2013).
- Ball, J. M., Lee, M. M., Hey, A. & Snaith, H. J. Low-temperature processed meso-superstructured to thin-film perovskite solar cells. *Energy Environ. Sci.* **6**, 1739–1743 (2013).
- Bi, D., Yang, L., Boschloo, G., Hagfeldt, A. & Johansson, E. M. J. Effect of different hole transport materials on recombination in CH₃NH₃PbI₃ perovskite-sensitized mesoscopic solar cells. *J. Phys. Chem. Lett.* **4**, 1532–1536 (2013).
- Heo, J. H. *et al.* Efficient inorganic–organic hybrid heterojunction solar cells containing perovskite compound and polymeric hole conductors. *Nature Photon.* **7**, 486–492 (2013).
- Beckmann, A. A review of polytypism in lead iodide. *Cryst. Res. Technol.* **45**, 455–460 (2010).
- Baikie, T. *et al.* Synthesis and crystal chemistry of the hybrid perovskite (CH₃NH₃)PbI₃ for solid-state sensitized solar cell applications. *J. Mater. Chem. A* **1**, 5628–5641 (2013).
- Liang, K., Mitzi, D. B. & Prikas, M. T. Synthesis and characterization of organic-inorganic perovskite thin films prepared using a versatile two-step dipping technique. *Chem. Mater.* **10**, 403–411 (1998).
- Beberwyck, B. J. & Alivisatos, A. P. Ion exchange synthesis of III–V nanocrystals. *J. Am. Chem. Soc.* **134**, 19977–19980 (2012).
- Luther, J. M., Zheng, H., Sadtler, B. & Alivisatos, A. P. Synthesis of PbS nanorods and other ionic nanocrystals of complex morphology by sequential cation exchange reactions. *J. Am. Chem. Soc.* **131**, 16851–16857 (2009).
- Li, H. *et al.* Sequential cation exchange in nanocrystals: preservation of crystal phase and formation of metastable phases. *Nano Lett.* **11**, 4964–4970 (2011).
- Gurina, G. I. & Savchenko, K. V. Intercalation and formation of complexes in the system of lead(II) iodide–ammonia. *J. Solid State Chem.* **177**, 909–915 (2004).
- Preda, N., Mihut, L., Baibarac, M., Baltog, I. & Lefrant, S. A distinctive signature in the Raman and photoluminescence spectra of intercalated PbI₂. *J. Phys. Condens. Matter* **18**, 8899–8912 (2006).
- Warren, R. F. & Liang, W. Y. Raman spectroscopy of new lead iodide intercalation compounds. *J. Phys. Condens. Matter* **5**, 6407–6418 (1993).
- Burschka, J. *et al.* Tris(2-(1H-pyrazol-1-yl)pyridine)cobalt(III) as p-type dopant for organic semiconductors and its application in highly efficient solid-state dye-sensitized solar cells. *J. Am. Chem. Soc.* **133**, 18042–18045 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Marszalek for recording SEM images, M. Tschumi for help with the stability measurements and K. Schenk for the XRD characterization. We acknowledge financial support from Aisin Cosmos R&D Co., Ltd. Japan; the European Community's Seventh Framework Programme (FP7/2007–2013) ENERGY.2012.10.2.1; NANOMATCELL, grant agreement no. 308997; the Global Research Laboratory Program, Korea; the Center for Advanced Molecular Photovoltaics (award no. KUS-C1-015-21) of King Abdullah University of Science and Technology; and Solvay S.A. M.K.N. thanks the World Class University programmes (Photovoltaic Materials, Department of Material Chemistry, Korea University), funded by the Ministry of Education, Science and Technology through the National Research Foundation of Korea (R31-2008-000-10035-0). M.G. thanks the Max Planck Society for a Max Planck Fellowship at the MPI for Solid State Research in Stuttgart, Germany; the King Abdulaziz University, Jeddah and the Nanyang Technological University, Singapore for Adjunct Professor appointments; and the European Research Council for an Advanced Research Grant (ARG 247404) funded under the "Mesolight" project.

Author Contributions J.B. developed the basic concept, carried out the spectroscopic characterization, fabricated and characterized photovoltaic devices, and coordinated the project. N.P. fabricated and characterized photovoltaic devices, optimized device performance and fabricated high-performance devices for the certification. S.-J.M. contributed to the fabrication and characterization of photovoltaic devices. R.H.-B. contributed to the spectroscopic characterization and data analysis. P.G. synthesized CH₃NH₃I. M.G. and J.B. analysed the data and wrote the paper. M.K.N. contributed to the supervision of the project. M.G. had the idea for, and directed, the project. All authors reviewed the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G. (michael.graetzel@epfl.ch).

METHODS

Materials. Unless stated otherwise, all materials were purchased from Sigma-Aldrich or Acros Organics and used as received. Spiro-MeOTAD was purchased from Merck KGaA. $\text{CH}_3\text{NH}_3\text{I}$ was synthesized according to a reported procedure³.

Device fabrication. First, laser-patterned, FTO-coated glass substrates (Tec15, Pilkington) were cleaned by ultrasonication in an alkaline, aqueous washing solution, rinsed with deionized water, ethanol and acetone, and subjected to an O_3 /ultraviolet treatment for 30 min. A 20–40-nm-thick TiO_2 compact layer was then deposited on the substrates by aerosol spray pyrolysis at 450 °C using a commercial titanium diisopropoxide bis(acetylacetonate) solution (75% in 2-propanol, Sigma-Aldrich) diluted in ethanol (1:39, volume ratio) as precursor and oxygen as carrier gas. After cooling to room temperature (~ 25 °C), the substrates were treated in an 0.04 M aqueous solution of TiCl_4 for 30 min at 70 °C, rinsed with deionized water and dried at 500 °C for 20 min.

The mesoporous TiO_2 layer composed of 20-nm-sized particles was deposited by spin coating at 5,000 r.p.m. for 30 s using a commercial TiO_2 paste (Dyesol 18NRT, Dyesol) diluted in ethanol (2:7, weight ratio). After drying at 125 °C, the TiO_2 films were gradually heated to 500 °C, baked at this temperature for 15 min and cooled to room temperature. Prior to their use, the films were again dried at 500 °C for 30 min.

PbI_2 was dissolved in *N,N*-dimethylformamide at a concentration of 462 mg ml^{-1} ($\sim 1 \text{ M}$) under stirring at 70 °C. The solution was kept at 70 °C during the whole procedure. The mesoporous TiO_2 films were then infiltrated with PbI_2 by spin coating at 6,500 r.p.m. for 90 s and dried at 70 °C for 30 min. After cooling to room temperature, the films were dipped in a solution of $\text{CH}_3\text{NH}_3\text{I}$ in 2-propanol (10 mg ml^{-1}) for 20 s, rinsed with 2-propanol and dried at 70 °C for 30 min.

The HTM was then deposited by spin coating at 4,000 r.p.m. for 30 s. The spin-coating formulation was prepared by dissolving 72.3 mg (2,2',7,7'-tetrakis(*N*,*N*-di-*p*-methoxyphenylamine)-9,9-spirobifluorene) (spiro-MeOTAD), 28.8 μl 4-tert-butylpyridine, 17.5 μl of a stock solution of 520 mg ml^{-1} lithium bis(trifluoromethylsulphonyl)imide in acetonitrile and 29 μl of a stock solution of 300 mg ml^{-1} tris(2-(1H-pyrazol-1-yl)-4-tert-butylpyridine)cobalt(III) bis(trifluoromethylsulphonyl)imide in acetonitrile in 1 ml chlorobenzene.

Finally, 80 nm of gold was thermally evaporated on top of the device to form the back contact. The device fabrication was carried out under controlled atmospheric conditions and a humidity of <1%. For the fabrication of the best-performing devices exhibiting a PCE of 15%, slightly modified conditions were used. First, PbI_2 was spin-cast at 6,500 r.p.m. for 5 s. Second, the samples were subjected to a 'pre-wetting' by dipping in 2-propanol for 1–2 s before being dipped in the solution of $\text{CH}_3\text{NH}_3\text{I}$ and 2-propanol.

Device characterization. Current–voltage characteristics were recorded by applying an external potential bias to the cell while recording the generated photocurrent with a digital source meter (Keithley Model 2400). The light source was a 450-W xenon lamp (Oriel) equipped with a Schott K113 Tempax sunlight filter (Prazisions Glas & Optik GmbH) to match the emission spectrum of the lamp to the AM1.5G standard. Before each measurement, the exact light intensity was determined using a calibrated Si reference diode equipped with an infrared cut-off filter (KG-3, Schott). IPCE spectra were recorded as functions of wavelength under a constant white light bias of approximately 5 mW cm^{-2} supplied by an array of white light-emitting diodes. The excitation beam coming from a 300-W xenon lamp (ILC Technology) was focused through a Gemini-180 double monochromator (Jobin Yvon Ltd) and chopped at approximately 2 Hz. The signal was recorded using a Model SR830 DSP Lock-In Amplifier (Stanford Research Systems). All measurements were conducted using a non-reflective metal aperture of 0.285 cm^2 to define the active area of the device and avoid light scattering through the sides.

Long-term stability. For long-term stability tests, the devices were sealed in argon using a 50- μm -thick hot-melting polymer and a microscope coverslip, and subjected to constant light soaking at approximately 100 mW cm^{-2} . The light source was an array of white light-emitting diodes (LXM3-PW51 4000K, Philips). During the testing, the devices were maintained at their maximum power point using electronic control and at a temperature of about 45 °C. An automatic *J*–*V* measurement at different light intensities (0%, 1%, 10%, 50% and 100% the solar value) was made every 2 h.

Optical spectroscopy. Mesoporous TiO_2 films were deposited on microscope glass slides and infiltrated with PbI_2 following the above-mentioned procedure. The samples were then placed vertically in a standard cuvette of 10-mm path length using a Teflon holder. A solution of $\text{CH}_3\text{NH}_3\text{I}$ in 2-propanol was then rapidly injected into the cuvette while either the photoluminescence or the optical transmission was monitored. Photoluminescence measurements were carried out on a Horiba Jobin Yvon Fluorolog spectrofluorometer. Optical absorption measurements were carried out on a Varian Cary 5 spectrophotometer.

X-ray diffraction measurements. For XRD measurements, flat PbI_2 and $\text{TiO}_2/\text{PbI}_2$ nanocomposites were deposited on microscope glass slides using the above-mentioned procedures. X-ray powder diagrams were recorded on an X'Pert MPD PRO from PANalytical equipped with a ceramic tube (Cu anode, $\lambda = 1.54060 \text{ \AA}$), a secondary graphite (002) monochromator and a RTMS X'Celerator detector, and operated in BRAGG-BRENTANO geometry. The samples were mounted without further modification, and the automatic divergence slit and beam mask were adjusted to the dimensions of the thin films. A step size of 0.008 deg was chosen and an acquisition time of up to 7.5 min deg^{-1} . A baseline correction was applied to all X-ray powder diagrams to remove the broad diffraction peak arising from the amorphous glass slide.

Myomaker is a membrane activator of myoblast fusion and muscle formation

Douglas P. Millay¹, Jason R. O'Rourke¹, Lillian B. Sutherland¹, Svetlana Bezprozvannaya¹, John M. Shelton², Rhonda Bassel-Duby¹ & Eric N. Olson¹

Fusion of myoblasts is essential for the formation of multi-nucleated muscle fibres. However, the identity of muscle-specific proteins that directly govern this fusion process in mammals has remained elusive. Here we identify a muscle-specific membrane protein, named myomaker, that controls myoblast fusion. Myomaker is expressed on the cell surface of myoblasts during fusion and is downregulated thereafter. Overexpression of myomaker in myoblasts markedly enhances fusion, and genetic disruption of myomaker in mice causes perinatal death due to an absence of multi-nucleated muscle fibres. Remarkably, forced expression of myomaker in fibroblasts promotes fusion with myoblasts, demonstrating the direct participation of this protein in the fusion process. Pharmacological perturbation of the actin cytoskeleton abolishes the activity of myomaker, consistent with previous studies implicating actin dynamics in myoblast fusion. These findings reveal a long-sought myogenic fusion protein that controls mammalian myoblast fusion and provide new insights into the molecular underpinnings of muscle formation.

Myoblast fusion is a complex and tightly controlled process required for the formation of skeletal muscle fibres¹. The fusion process must be highly cell-type specific to ensure that fusogenic myoblasts do not form syncytia with non-muscle cell types. Although the transcriptional mechanisms governing skeletal muscle development have been explained in detail^{2–5}, the mechanisms that coordinate myoblast fusion remain incompletely understood, and no muscle-specific protein that directly regulates myoblast fusion has been identified in mammals^{6,7}. In contrast, numerous proteins involved in cell–cell adhesion and actin dynamics have been implicated in myoblast fusion^{8–12}. However, none of these proteins is muscle-specific, necessary and sufficient for mammalian myoblast fusion, suggesting that muscle-specific components of this process remain to be discovered. Here we describe a muscle-specific membrane protein called myomaker that is transiently expressed during myogenesis and is both necessary and sufficient to promote myoblast fusion *in vivo* and *in vitro*.

Discovery and regulation of myomaker

To search for new skeletal-muscle regulatory genes, we interrogated the National Center for Biotechnology Information UniGene database for genes of unknown function with expression profiles similar to those of MyoD (*Myod*) and myogenin (*Myog*), which encode important muscle-specific transcription factors^{13,14}. Among the genes identified in this screen was transmembrane protein 8c (*Tmem8c*), which had not been previously studied. On the basis of the observations described below, we named this gene myomaker.

During mouse embryogenesis, the myomaker gene is robustly expressed in the myotomal compartment of the somites, and later is expressed in limb buds and axial skeletal muscles (Fig. 1a and Supplementary Fig. 1a). Expression of the myomaker gene in the myotomes coincides with expression of other known muscle transcripts, such as *Myog* and M-cadherin (Supplementary Fig. 1a). Myomaker gene messenger RNA (mRNA) is expressed in developing skeletal muscle and is subsequently downregulated upon completion of muscle formation, similar to the expression pattern of *Myod* and *Myog* (Fig. 1b). Myomaker gene

expression was not detected in tissues other than skeletal muscle in embryonic day (E)19 embryos (Supplementary Fig. 1b, c). In the C2C12 skeletal muscle cell line, the myomaker gene mimics *Myog* expression, increasing sharply during differentiation and fusion (Fig. 1c).

To begin to assess the function of myomaker in skeletal muscle, we obtained embryonic stem cells that contained a LacZ-Neo cassette in intron 1 of the myomaker locus (Supplementary Fig. 2a). In this allele, exon 1 of the myomaker gene is spliced to *lacZ*, preventing expression of a functional myomaker gene transcript. We refer to mice heterozygous and homozygous for the myomaker-*lacZ* allele as myomaker^{+/-} and myomaker^{-/-} mice, respectively. 5-Bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) staining of myomaker^{+/-} mice showed expression of the targeted *lacZ* allele specifically in skeletal muscle, and not in other muscle tissues or non-muscle tissues (Fig. 1d and Supplementary Fig. 2b, c). Like the endogenous myomaker gene, skeletal muscle expression of the myomaker-*lacZ* allele declined postnatally (Supplementary Fig. 2d).

Adult skeletal muscle regenerates in response to damage, owing to the activation of satellite cells, which fuse with residual muscle fibres^{4,5}. We tested whether myomaker gene expression is re-activated during adult muscle regeneration by inducing muscle injury in adult mice. Expression of the myomaker-*LacZ* allele and myomaker gene mRNA was strongly induced in regenerating muscle after cardiotoxin injury (Fig. 1e and Supplementary Fig. 2e). We conclude that the myomaker gene is expressed specifically in skeletal muscle during embryogenesis and adult muscle regeneration.

Myomaker is required for muscle formation

We generated myomaker^{-/-} mice by interbreeding of heterozygous mice. Myomaker transcripts were absent in skeletal muscle of myomaker^{-/-} mice, confirming that the targeting strategy created a null allele (Supplementary Fig. 2f). Myomaker^{-/-} mice were observed at normal Mendelian ratios at E15 and E17.5; however, we failed to detect any live myomaker^{-/-} mice at postnatal day 7, suggesting earlier lethality due to muscle dysfunction (Supplementary Fig. 2g).

¹Department of Molecular Biology, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. ²Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.

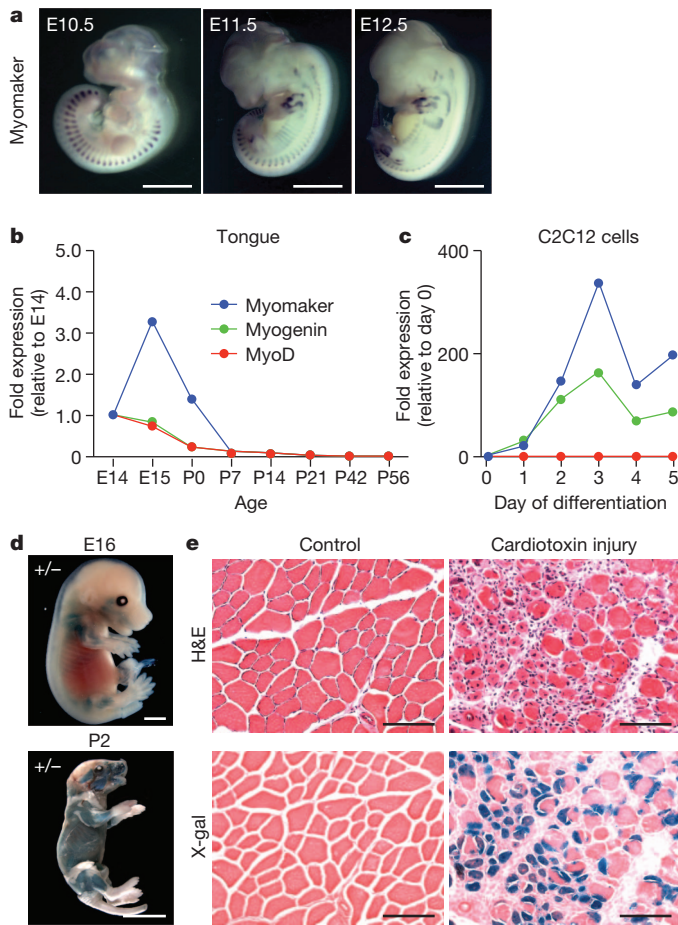


Figure 1 | Muscle-specific expression of myomaker. **a**, *In situ* hybridization for myomaker in WT embryos illustrates muscle specificity. **b**, Quantitative PCR for the myomaker gene, *Myog* and *MyoD* on tongues at the indicated ages shows downregulation after myogenesis. **c**, Gene expression during differentiation of C2C12 myoblasts. The key in **b** also applies to **c**. **d**, X-gal staining on E16 and postnatal day 2 myomaker^{+/-} (myomaker-*LacZ*) mice confirms expression in all skeletal muscles. **e**, Cardiotoxin-injured and X-gal-stained tibialis anterior muscle from 6-week-old myomaker^{+/-} mice shows the re-activation of myomaker. Serial sections stained with haematoxylin and eosin (H&E) indicate muscle injury. Control represents uninjured contralateral tibialis anterior muscle. Scale bars: **a**, **d**, 2 mm; **e**, 200 μ m.

Full-term myomaker^{-/-} embryos were alive, as their hearts were beating, but were paralysed and kyphotic with flaccid limbs, hallmarks of skeletal muscle deficiency (Fig. 2a). Strikingly, no semblance of differentiated muscle tissue was present in the trunk, limbs or head of myomaker^{-/-} animals (Fig. 2b and Supplementary Fig. 3a, b).

Muscle formation requires myoblast specification, migration, differentiation and fusion²⁻⁵. In principle, dysfunction of one or more of these processes could contribute to lethality and lack of muscle formation in myomaker^{-/-} embryos. To begin to define the mechanistic actions of myomaker, we tested the functionality of these processes. The muscle-specific transcription factors, MyoD and myogenin, were expressed normally in myomaker^{-/-} embryos (Supplementary Fig. 3c, d), suggesting that specification of the skeletal muscle lineage occurred normally in the absence of myomaker. Muscle tissues were present in myomaker^{-/-} embryos, indicating that muscle precursor cells were organized appropriately in the absence of myomaker (Supplementary Fig. 3e). Desmin, a marker of muscle cells, was expressed comparably in myomaker^{-/-} and wild-type (WT or +/+) forelimbs, confirming that myoblast migration was unaltered (Supplementary Fig. 3f). These findings suggested myomaker functions after myoblast specification and migration. Longitudinal sections through hindlimb

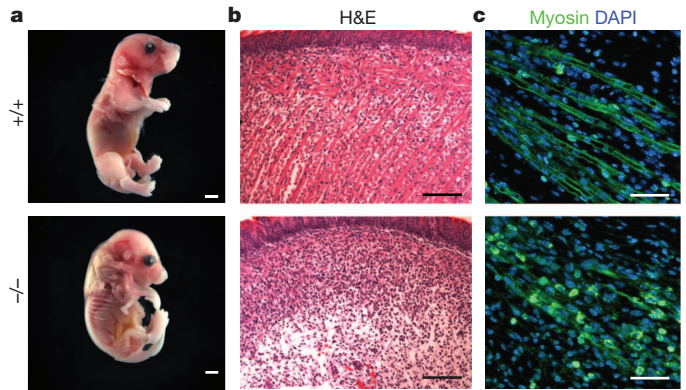


Figure 2 | Myomaker is essential for skeletal muscle development. **a**, Full-term WT (+/+) and myomaker^{-/-} embryos were dissected and skinned to illustrate the lack of muscle surrounding myomaker^{-/-} limbs. **b**, Paraffin sectioning and H&E staining on tongues show a lack of muscle fibres in E17.5 myomaker^{-/-} embryos. **c**, Longitudinal sections of E14 hindlimb muscles stained with a myosin antibody to determine the multi-nucleation of the muscle cells. WT limbs show myofibres containing several nuclei, which are absent in myomaker^{-/-} sections. Scale bars: **a**, 2 mm; **b**, 100 μ m; **c**, 40 μ m.

muscles of myomaker^{-/-} embryos at E14 showed the expression of myosin, a muscle differentiation marker, but an absence of multi-nucleated myofibres (Fig. 2c). These findings imply that myomaker^{-/-} myoblasts can activate muscle-specific gene expression and differentiate, but lack the ability to fuse.

Myomaker^{-/-} muscle tissues contained only mono-nucleated cells; however, the cell number was clearly reduced in each muscle analysed. One possible explanation for this decrease is cell death, which has previously been associated with a failure to fuse^{15,16}. Indeed, TdT-mediated dUTP nick end labelling (TUNEL) analysis showed increased apoptotic nuclei in muscle forming regions of myomaker^{-/-} mice, suggesting that fusion-defective myoblasts are non-viable (Supplementary Fig. 3g).

Myomaker controls myoblast fusion

To confirm definitively that myomaker functions in myoblast fusion, we used several *in vitro* differentiation assays using primary myoblasts and the C2C12 muscle cell line. First, we isolated myoblasts from WT and myomaker^{-/-} embryos. After 3 days of differentiation, WT myoblasts formed extensive myotubes containing many nuclei (Fig. 3a). In contrast, most myomaker^{-/-} myoblasts remained mono-nucleated, with only a small percentage forming bi-nucleated myosin positive cells (Fig. 3a, b and Supplementary Fig. 4a). Quantification of the differentiation index showed no differences in the ability of myomaker^{-/-} myoblasts to express myosin; however, the fusion index was markedly reduced compared with WT myoblasts (Fig. 3c and Supplementary Fig. 4b), even when plated for prolonged periods at higher density than WT myoblasts, indicating that fusion was blocked rather than simply delayed (Supplementary Fig. 4c). We conclude that the lack of muscle formation in myomaker^{-/-} embryos is due to a block of myoblast fusion, representing the cellular mechanism of myomaker function.

To test whether myomaker was a limiting factor in myoblast fusion, we infected C2C12 cells with a myomaker retrovirus 1 day before differentiation and assessed the consequences on myoblast fusion. Myomaker overexpression caused a marked increase in fusion after 4 days of differentiation (Fig. 3d). The kinetics of induction of myogenin and myosin, and maximal amounts of expression of the terminal differentiation genes (*Myog*, *Ckm* and *Myh4*), were comparable in myomaker-infected cells and cells infected with a green fluorescent protein (GFP) control virus (Supplementary Fig. 4d, e). Despite no differences in expression of muscle differentiation factors, we observed a robust increase in the appearance of myotubes with several nuclei in

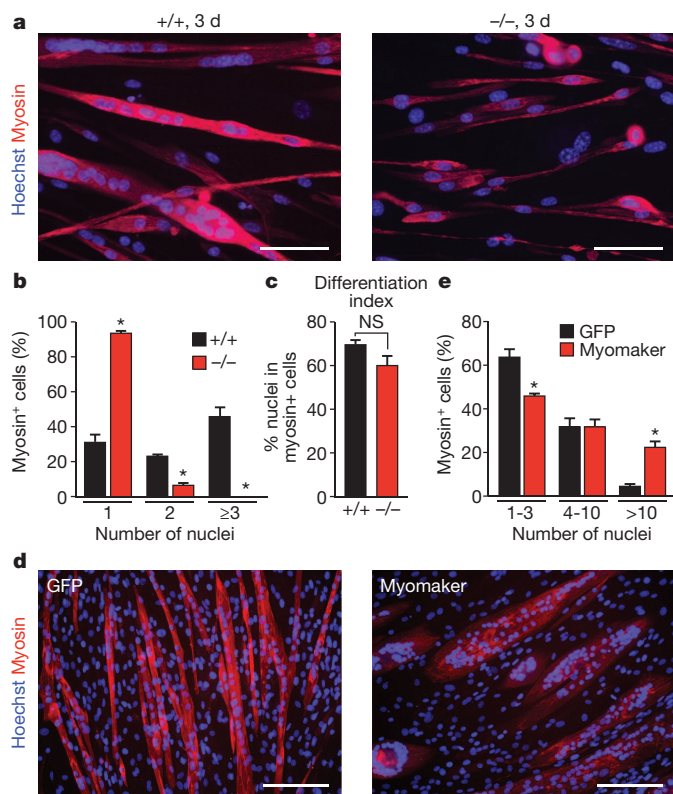


Figure 3 | Control of myoblast fusion by myomaker. **a**, Myoblasts from WT (+/+) and myomaker^{-/-} E17 embryos were differentiated for 3 days, and stained for myosin and a nuclear stain (Hoechst). Myomaker^{-/-} myoblasts failed to fuse. **b**, The number of nuclei present in a myosin⁺ cell indicates myomaker^{-/-} myoblasts cannot form myotubes with three or more nuclei. **c**, Differentiation index, calculated as the percentage of nuclei in myosin⁺ cells, indicated null myoblasts can activate the myogenic program. NS, not statistically significant. **d**, C2C12 cells infected with a retrovirus encoding GFP or myomaker were induced to differentiate for 4 days then stained with a myosin antibody and Hoechst (nuclei). **e**, The percentage of myosin⁺ cells that contained the indicated number of nuclei. Quantification was performed after 3 days of differentiation in **b**, **c**, and after 4 days in **e**. Scale bars: **a**, 100 μ m; **e**, 200 μ m. Data are presented as mean \pm s.e.m. from three independent experiments. * $P < 0.05$ compared with WT in **b** and **c**, or with GFP-infected cells in **e**.

the cultures infected with myomaker, further indicating that myomaker functions specifically in myoblast fusion and does not regulate differentiation per se (Supplementary Fig. 4f). Quantification of the fusion index and the number of nuclei per myotube indicated a robust activity of myomaker to increase the fusion capability of these cells (Fig. 3e and Supplementary Fig. 3g, h). Furthermore, through live cell imaging, we visualized myotube-myotube fusion in myomaker-infected cells (Supplementary Movie 1). These data demonstrate that myomaker is sufficient to enhance C2C12 myoblast fusion.

Myomaker is 221 amino acids in length and highly conserved across vertebrate organisms, ranging from fish to humans (Supplementary Fig. 5a). Analysis of the hydrophobicity of myomaker using a Kyte–Doolittle plot showed extensive regions of hydrophobic character, suggesting this protein may localize to a cellular membrane (Supplementary Fig. 5b). Myomaker does not contain predicted *N*-glycosylation sites. At the carboxy terminus, myomaker possesses a C-A-A-X motif, the consensus for isoprenylation, which mediates membrane association¹⁷. Myomaker shares limited homology to a family of putative transmembrane hydrolases, named the CREST family¹⁸, but it lacks a potentially critical histidine residue thought to be important for catalytic activity of hydrolases. The closest relative, Tmem8b, shares homology with myomaker/Tmem8c in three hydrophobic domains; however, Tmem8b is not muscle-specific and its

forced expression in C2C12 cells did not promote fusion (data not shown). There is also a related protein in *Drosophila*, but it is more similar to Tmem8a and Tmem8b than to myomaker/Tmem8c.

To analyse the cellular distribution of myomaker, we engineered a Flag epitope after amino acid 61, in a region of the protein that would not be predicted to perturb the hydrophobic domains (Supplementary Fig. 5b). The Flag-tagged myomaker protein, referred to as myomaker-Flag, was detected in whole-cell lysates, by Flag western blots (Supplementary Fig. 6a). Retroviral expression of myomaker-Flag in C2C12 cells confirmed that insertion of the Flag epitope did not alter the function of myomaker as assayed by its ability to enhance myoblast fusion (Supplementary Fig. 6b). Fractionation of C2C12 cells infected with myomaker-Flag into membrane and cytosolic fractions, showed exclusive localization to the membrane fraction (Supplementary Fig. 6c). Myomaker-Flag was readily detected on the surface of myoblasts, by staining live cells with a Flag antibody, a common method used to detect plasma membrane proteins¹⁹ (Fig. 4a). Moreover, in myoblast cultures undergoing fusion, myomaker-Flag was detected at sites of cell–cell interaction (Fig. 4b). Immunocytochemistry of fixed and permeabilized C2C12 cells expressing myomaker-Flag showed intracellular vesicle localization of myomaker-Flag, as expected for a membrane protein (Fig. 4c). Co-staining with intracellular organelle markers showed some co-localization with endosomes and endoplasmic reticulum (Supplementary Fig. 6d), suggesting that myomaker transits through one or more intracellular membrane compartments.

Myoblast fusion requires actin–cytoskeletal reorganization^{15,16,20–23}. Treating C2C12 cells with cytochalasin D and lantrunculin B, which perturb the cytoskeleton, completely blocked fusion in cells infected with GFP or myomaker virus, suggesting that actin nucleation is required for the fusogenic function of myomaker (Supplementary Fig. 7a). After treatment with cytochalasin D, myomaker-Flag was properly localized to the membrane, indicating that actin dynamics do not regulate transport of the protein to the cell surface (Supplementary Fig. 7b).

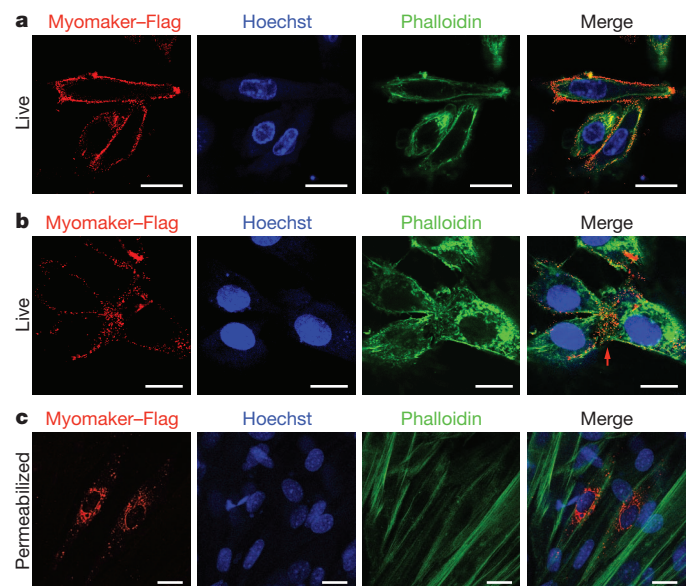


Figure 4 | Myomaker is expressed on the cell membrane of myoblasts. **a**, C2C12 cells were infected with myomaker-Flag and live cells were stained 2 days after differentiation with Flag antibody on ice. After Flag staining, cells were then fixed, permeabilized and stained with Phalloidin (F-actin) and Hoechst (nuclei) to illustrate cell membrane localization of myomaker-Flag. **b**, Cells were stained as in **a** to visualize myomaker-Flag in fusing cultures. The red arrow depicts sites of cell–cell interaction. **c**, Myomaker-Flag-infected C2C12 cells were fixed, permeabilized and stained with Flag antibody, Phalloidin and Hoechst to show the vesicle localization of the intracellular protein. Scale bars, 20 μ m.

Investigation of the fusogenic functions of myomaker

To understand the mechanism of action of myomaker further, we performed cell-mixing experiments using primary myoblasts from WT, myomaker^{+/-} and myomaker^{-/-} embryos (Fig. 5a). After differentiation for 4 days, we visualized β -galactosidase expression from the lacZ allele in myomaker^{+/-} and myomaker^{-/-} myoblasts to monitor fusion between different myoblast populations. As a co-stain, we used Nuclear Fast Red, which stains a nucleus red and confers a pink appearance in the cytoplasm of cells. Myomaker^{+/-} myoblasts formed multi-nucleated myotubes alone, without WT myoblasts, whereas myomaker^{-/-} myoblasts failed to fuse (Fig. 5a). Chimaeric myotubes (blue/pink) were apparent in cultures containing WT and myomaker^{+/-} myoblasts, indicating fusion between these two myoblast populations (Fig. 5a). In cultures containing both WT and myomaker^{-/-} myoblasts, we observed myotubes containing LacZ staining emanating from myomaker^{-/-} myoblasts (Fig. 5a). Quantification of the percentage of LacZ⁺ myotubes with three or more nuclei showed that myomaker^{-/-} myoblasts could only form these structures in the presence of WT myoblasts (Fig. 5b). We conclude that a cell with a functional copy of the myomaker gene can fuse with a myomaker^{-/-} myoblast, suggesting that myomaker is absolutely required on the surface of only one of the fusing muscle cells. We further investigated this possibility by analysing expression of myomaker-Flag in C2C12 cells and detected myomaker-Flag in mono-nuclear C2C12 cells but not in previously fused multi-nucleated myotubes (Fig. 5c).

To determine whether overexpression of myomaker could allow fusion of fibroblasts, a cell type that lacks fusion capability, we infected 10T1/2 fibroblasts with a GFP virus and either empty virus, as a control, or myomaker virus and then mixed these fibroblasts with C2C12 cells (Fig. 5d). We did not detect fusion of GFP-empty-virus-infected fibroblasts with myosin-positive cells; however, GFP-myomaker-infected fibroblasts robustly fused with C2C12 cells (Fig. 5d and Supplementary Movie 2). Quantification of the myotubes expressing both GFP and myosin confirmed a striking ability of fibroblasts expressing myomaker to fuse with myoblasts (Fig. 5e).

To control for the possibility that myomaker-expressing fibroblasts were leaky and allowed GFP to diffuse into C2C12 myotubes, we designed a complementary cell mixing experiment in which we tracked fibroblast nuclei by labelling with BrdU, followed by mixing with dsRed-infected C2C12 cells (Supplementary Fig. 8a). BrdU-positive nuclei from fibroblasts expressing myomaker were detected within C2C12 myotubes, confirming that myomaker expression was sufficient to direct the fusion of fibroblasts to myoblasts (Supplementary Fig. 8a, b). Myomaker was not sufficient to induce fusion of fibroblasts in the absence of myoblasts. The finding that myomaker can promote fusion of fibroblasts to myoblasts but cannot promote fibroblast-fibroblast fusion suggests that additional myoblast cell-surface proteins are required for proper fusogenic engagement of the two membranes.

Discussion

There are several types of membrane fusion, including virus-cell fusion, intracellular vesicle fusion and cell-cell fusion¹. Similarities exist between different fusion mechanisms, but relatively little is known about cell-cell fusion compared with other fusion processes, or the identities of fusogenic proteins that directly merge intercellular membranes. Our findings identify myomaker as a muscle-specific plasma membrane protein expressed specifically during times of myoblast fusion and required for the formation of multinucleated myofibres. Although surface glycoproteins, including cadherins, β -1 integrin, MOR23 and Adam12 (refs 8–12), have been shown to influence myoblast fusion, myomaker, to our knowledge, is the only muscle-specific protein yet identified that is essential for myoblast fusion in mammals. The absence of multinucleated myofibres in myomaker^{-/-} mice demonstrates the requirement of this membrane protein for the formation of all skeletal muscles.

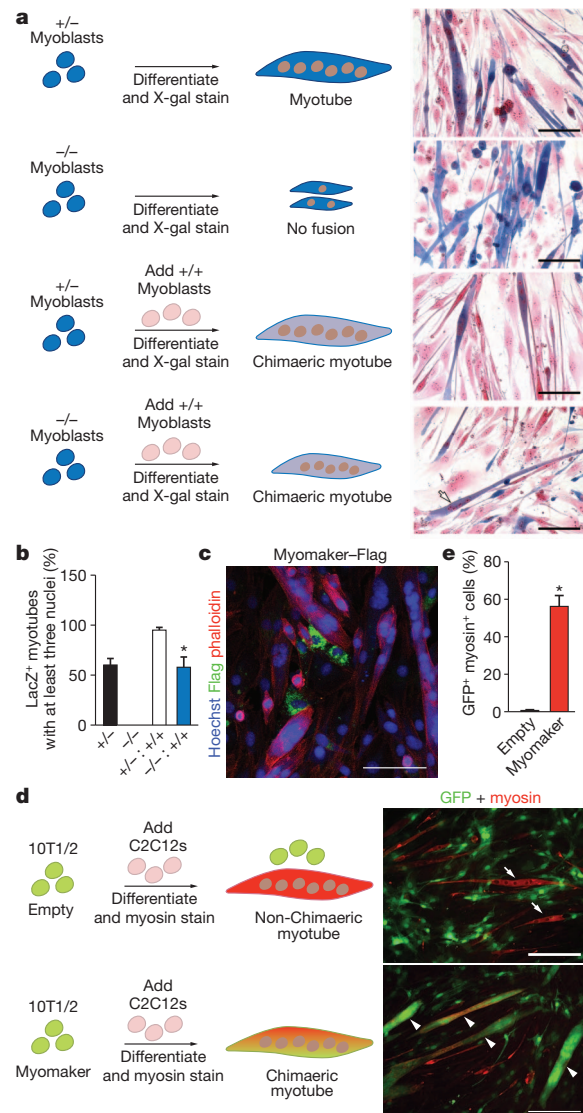


Figure 5 | Myomaker participates in the myoblast membrane fusion reaction. **a**, Myomaker^{+/-} and myomaker^{-/-} myoblasts express LacZ, and were either plated alone or mixed with WT myoblasts, induced to differentiate for 4 days and stained with X-gal and Nuclear Fast Red to determine the amount of fusion. Myomaker^{+/-} myoblasts, alone or in the presence of WT myoblasts, fused normally, illustrated by myotubes with robust LacZ staining. Myomaker^{-/-} myoblasts alone did not fuse. Addition of WT myoblasts to myomaker^{-/-} myoblasts resulted in chimaeric myotubes (arrow), indicating fusion between the two cell populations. **b**, The percentage of LacZ⁺ myotubes containing at least three nuclei shows null myoblasts can only form myotubes with three or more nuclei in the presence of WT myoblasts. **c**, Phalloidin and Flag staining of C2C12 myoblasts after infection with myomaker-Flag illustrates the lack of Flag staining in myotubes. **d**, 10T1/2 fibroblasts were infected with GFP-retrovirus and either Empty- or myomaker-retrovirus, then mixed with C2C12 cells and differentiated. Myotube formation was monitored by myosin staining, and fusion of fibroblasts was determined by visualization of GFP in myosin⁺ myotubes. Myosin⁺ GFP⁺ myotubes (arrowheads) are evident in cultures containing myomaker-infected fibroblasts, whereas myosin⁺ GFP⁻ myotubes (arrows) were observed in empty-infected cultures. **e**, The percentage of GFP⁺ fibroblasts, infected with empty- or myomaker-retrovirus, that fused to myosin⁺ myoblasts. Scale bars: **a**, 100 μ m; **d**, 200 μ m. Data are presented as mean \pm s.e.m. from three independent experiments. **P* < 0.05 compared with WT in **b** and with empty in **e**.

Myoblast fusion is a multistep process requiring intimate cell-cell interaction followed by membrane coalescence accompanied by actin-cytoskeletal dynamics that drive cell merger. Myomaker clearly participates in the membrane fusion reaction, as demonstrated by its

ability to stimulate myoblast fusion and the fusion of fibroblasts to myoblasts. The inability of myomaker alone to induce fusion of fibroblasts suggests it may require activation or other myoblast proteins to exert its fusogenic activity, probably reflecting a requirement for close membrane apposition to allow membrane merger. Further evidence that other myoblast proteins are required for fusion is our finding that WT myoblasts can fuse with myomaker^{-/-} myoblasts. The requirement for interactions between membrane proteins on opposite cells during myoblast fusion has been shown in zebrafish and *Drosophila*^{6,24}, suggesting the molecular regulation of myoblast fusion differs from that of virus–cell fusion, which mainly requires the expression of a fusogenic protein²⁵. Changes in the actin cytoskeleton are required for cell–cell fusion^{26,27}. Consistent with this, the activity of myomaker is abolished by cytochalasin D and latrunculin B, which disrupt cytoskeletal events required for fusion, indicating that myomaker depends on the cytoskeleton to exert its function. The discovery of myomaker as a potent myoblast fusion protein opens new opportunities to dissect this fundamental cellular process at a molecular level and to understand how myoblast fusion is perturbed during muscle disease. Moreover, the ability of myomaker to drive fusion of non-muscle cells with muscle cells represents an interesting strategy for enhancing muscle repair.

METHODS SUMMARY

All experimental procedures with animals were approved by the Institutional Animal Care and Use Committee at University of Texas Southwestern Medical Center.

The myomaker/Tmem8c mouse strain used for this research project was created from embryonic stem cell Tmem8c clone (EPD0626_5_C12) obtained from the KOMP Repository (www.KOMP.org) and generated by the Wellcome Trust Sanger Institute²⁸.

Whole-mount *in situ* hybridization was performed as previously described²⁹. To isolate muscle progenitors, embryonic limbs were dissected and dissociated in 0.05% Collagenase D (Roche) in PBS at 37 °C for 2–3 h and plated on laminin-coated culture dishes.

Retroviral plasmid DNA was generated by subcloning myomaker and Flag-tagged myomaker complementary DNAs (cDNAs) into the retroviral vector pBabe-X³⁰. GFP and dsRed retrovirus have been described previously³¹.

WT myoblasts were mixed with either myomaker^{+/-} or myomaker^{-/-} myoblasts in equal ratios, plated on a well of a laminin-coated 12-well plate, and induced to differentiate the next day. 10T1/2 fibroblasts were infected with either GFP- and empty-retrovirus or GFP- and myomaker-retrovirus for 18 h, then mixed with C2C12 cells at a 1:1 ratio.

Full Methods and any associated references are available in the online version of the paper.

Received 12 February; accepted 3 June 2013.

- Chen, E. H. & Olson, E. N. Unveiling the mechanisms of cell–cell fusion. *Science* **308**, 369–373 (2005).
- Bentzinger, C. F., Wang, Y. X. & Rudnicki, M. A. Building muscle: molecular regulation of myogenesis. *Cold Spring Harb. Perspect. Biol.* **4**, a008342 (2012).
- Berkes, C. A. & Tapscott, S. J. MyoD and the transcriptional control of myogenesis. *Semin. Cell Dev. Biol.* **16**, 585–595 (2005).
- Buckingham, M. Myogenic progenitor cells and skeletal myogenesis in vertebrates. *Curr. Opin. Genet. Dev.* **16**, 525–532 (2006).
- Kang, J. S. & Krauss, R. S. Muscle stem cells in developmental and regenerative myogenesis. *Curr. Opin. Clin. Nutr. Metab. Care* **13**, 243–248 (2010).
- Abmayr, S. M. & Pavlath, G. K. Myoblast fusion: lessons from flies and mice. *Development* **139**, 641–656 (2012).
- Rochlin, K., Yu, S., Roy, S. & Baylies, M. K. Myoblast fusion: when it takes more to make one. *Dev. Biol.* **341**, 66–83 (2010).
- Charrasse, S. *et al.* M-cadherin activates Rac1 GTPase through the Rho-GEF trio during myoblast fusion. *Mol. Biol. Cell* **18**, 1734–1743 (2007).

- Charrasse, S., Meriane, M., Comunale, F., Blangy, A. & Gauthier-Rouviere, C. N-cadherin-dependent cell–cell contact regulates Rho GTPases and beta-catenin localization in mouse C2C12 myoblasts. *J. Cell Biol.* **158**, 953–965 (2002).
- Schwander, M. *et al.* β 1 integrins regulate myoblast fusion and sarcomere assembly. *Dev. Cell* **4**, 673–685 (2003).
- Griffin, C. A., Kafadar, K. A. & Pavlath, G. K. MOR23 promotes muscle regeneration and regulates cell adhesion and migration. *Dev. Cell* **17**, 649–661 (2009).
- Yagami-Hromasa, T. *et al.* A metalloprotease-disintegrin participating in myoblast fusion. *Nature* **377**, 652–656 (1995).
- Davis, R. L., Weintraub, H. & Lassar, A. B. Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell* **51**, 987–1000 (1987).
- Hasty, P. *et al.* Muscle deficiency and neonatal death in mice with a targeted mutation in the *myogenin* gene. *Nature* **364**, 501–506 (1993).
- Vasyutina, E., Martarelli, B., Brakebusch, C., Wende, H. & Birchmeier, C. The small G-proteins Rac1 and Cdc42 are essential for myoblast fusion in the mouse. *Proc. Natl Acad. Sci. USA* **106**, 8935–8940 (2009).
- Gruenbaum-Cohen, Y. *et al.* The actin regulator N-WASP is required for muscle-cell fusion in mice. *Proc. Natl Acad. Sci. USA* **109**, 11211–11216 (2012).
- Wright, L. P. & Philips, M. R. Thematic review series: lipid posttranslational modifications. CAAX modification and membrane targeting of Ras. *J. Lipid Res.* **47**, 883–891 (2006).
- Pei, J., Millay, D. P., Olson, E. N. & Grishin, N. V. CREST—a large and diverse superfamily of putative transmembrane hydrolases. *Biol. Direct* **6**, 37 (2011).
- Corcoran, J. A. & Duncan, R. Reptilian reovirus utilizes a small type III protein with an external myristylated amino terminus to mediate cell–cell fusion. *J. Virol.* **78**, 4342–4351 (2004).
- Chen, E. H. & Olson, E. N. Antisocial, an intracellular adaptor protein, is required for myoblast fusion in *Drosophila*. *Dev. Cell* **1**, 705–715 (2001).
- Chen, E. H., Pryce, B. A., Tzeng, J. A., Gonzalez, G. A. & Olson, E. N. Control of myoblast fusion by a guanine nucleotide exchange factor, Ioner, and its effector ARF6. *Cell* **114**, 751–762 (2003).
- Nowak, S. J., Nahirney, P. C., Hadjantonakis, A. K. & Baylies, M. K. Nap1-mediated actin remodeling is essential for mammalian myoblast fusion. *J. Cell Sci.* **122**, 3282–3293 (2009).
- Laurin, M. *et al.* The atypical Rac activator Dock180 (Dock1) regulates myoblast fusion *in vivo*. *Proc. Natl Acad. Sci. USA* **105**, 15446–15451 (2008).
- Powell, G. T. & Wright, G. J. Jamb and junc are essential for vertebrate myocyte fusion. *PLoS Biol.* **9**, e1001216 (2011).
- Oren-Suissa, M. & Podbilewicz, B. Cell fusion during development. *Trends Cell Biol.* **17**, 537–546 (2007).
- Wilson, S. J. & Snell, W. J. Microvilli and cell–cell fusion during fertilization. *Trends Cell Biol.* **8**, 93–96 (1998).
- Shilagardi, K. *et al.* Actin-propelled invasive membrane protrusions promote fusogenic protein engagement during cell–cell fusion. *Science* **340**, 359–363 (2013).
- Skarnes, W. C. *et al.* A conditional knockout resource for the genome-wide study of mouse gene function. *Nature* **474**, 337–342 (2011).
- Hargrave, M. & Koopman, P. *In situ* hybridization of whole-mount embryos. *Methods Mol. Biol.* **123**, 279–289 (2000).
- Kitamura, T. *et al.* Efficient screening of retroviral cDNA expression libraries. *Proc. Natl Acad. Sci. USA* **92**, 9146–9150 (1995).
- Song, K. *et al.* Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* **485**, 599–604 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Cabrera for graphics, the Transgenic Technology Center at University of Texas Southwestern Medical Center for embryonic stem cell injections, J.A. Richardson for histology and the University of Texas Southwestern Live Cell Imaging Facility (K. Luby-Phelps and A. Budge) for microscopy and live cell-imaging assistance. We thank D. Rosenbaum, W. Snell, S. Schmid, J. Seemann and members of the Olson laboratory for scientific discussions. D.P.M. was financed by a National Institutes of Health National Research Service Award Fellowship (F32AR05948403). This work was supported by grants from the National Institutes of Health (HL-077439, HL-111665, HL093039 and U01-HL-100401) and the Robert A. Welch Foundation (grant 1-0025) (to E.N.O.).

Author Contributions D.P.M. and E.N.O. conceived the project and designed the experiments. D.P.M., J.R.O., L.B.S., S.B. and J.M.S. performed experiments. D.P.M. and R.B.-D. wrote the animal protocol. D.P.M. and E.N.O. analysed the data and prepared the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.N.O. (eric.olson@utsouthwestern.edu).

METHODS

Generation of myomaker^{-/-} mice. The myomaker mouse strain used for this research project was created from embryonic stem cell Tmem8c clone (EPD0626_5_C12) obtained from the KOMP Repository (www.KOMP.org) and generated by the Wellcome Trust Sanger Institute²⁹. This clone was injected into 3.5-day-old C57BL/6 blastocysts by the Transgenic Core Facility at University of Texas Southwestern Medical Center. High-percentage chimaeric male mice were bred to C57BL/6 females to achieve germline transmission of the targeted allele. Myomaker^{+/-} mice were intercrossed to generate myomaker^{-/-} mice. All experimental procedures involving animals in this study were reviewed and approved by the University of Texas Southwestern Medical Center's Institutional Animal Care and Use Committee.

Quantitative real-time PCR (qPCR). Total RNA was extracted from either mouse tissue or cultured cells with TRIzol (Invitrogen) and cDNA synthesized using Superscript III reverse transcriptase with random hexamer primers (Invitrogen). Gene expression was assessed using standard qPCR approaches with either Power SYBR Green or Taqman Master Mix (Applied Biosystems). We used a 7900HT Fast Real-Time PCR machine (Applied Biosystems) with the following Sybr primers: myomaker-F: 5'-ATCGCTACCAAGAGGCGTT-3', myomaker-R: 5'-CAGCAGACAGAAACCAGG-3'. Taqman probes for myogenin, MyoD, Ckm and Myh4 were purchased from Applied Biosystems. Expression amounts were normalized to 18S and represented as fold change.

In situ hybridizations. For whole-mount *in situ* hybridization, embryos were fixed overnight in 4% PFA/PBS at 4 °C, then dehydrated in increasing concentrations of methanol and bleached with 6% H₂O₂/methanol for 1 h. Embryos were subsequently rehydrated, treated with proteinase K and fixed in 4% PFA, 0.2% glutaraldehyde for 20 min. Pre-hybridization (50% formamide, 5× SSC pH 4.5, 2% SDS, 2% blocking reagent (Roche), 250 µg ml⁻¹ transfer RNA, 100 µg ml⁻¹ heparin) was achieved at 70 °C for 1 h followed by incubation with digoxigenin-labelled probe overnight. Embryos were first washed with solution 1 (50% formamide, 2× SSC pH 4.5 and 1% SDS) three times, six times in solution 2 (100 mM Maleic Acid, 150 mM NaCl, 0.1% Tween-20, pH 7.5), then blocked with consecutive 1 h incubations with 2% blocking reagent/solution 2 and 2% blocking reagent/20% heat-inactivated goat serum/solution 2. To detect bound probe, we performed immunohistochemistry with anti-digoxigenin-alkaline phosphatase antibody (1:2,000, Roche). To develop the AP signal, embryos were washed with solution 1, then incubated with solution 4 (100 mM NaCl, 100 mM Tris-Cl, pH 9.5, 50 mM MgCl₂, 0.1% Tween-20) with developing reagents (0.25 mg ml⁻¹ NBT (nitro blue tetrazolium chloride) and 0.125 mg ml⁻¹ BCIP (5-bromo-4-chloro-3-indolyl phosphate, toluidine salt, Roche). Lastly, the embryos were washed with solution 4, fixed in 4% PFA/PBS at 4 °C overnight and imaged with a Zeiss 11 stereoscope. Full-length coding sequence was used to generate probes for both MyoD and myomaker by using the digoxigenin labelling kit (Roche) followed by purification with MicroSpin G-25 columns (Amersham).

Radioisotopic *in situ* hybridization was performed as previously described³². Briefly, sections were deparaffinized, permeabilized and acetylated before hybridization at 55 °C with riboprobes diluted in a mixture containing 50% formamide, 0.3 M NaCl, 20 mM Tris-HCl, pH 8.0, 5 mM EDTA, pH 8.0, 10 mM NaPO₄, pH 8.0, 10% dextran sulphate, 1× Denhardt's solution and 0.5 mg ml⁻¹ transfer RNA. After hybridization, the sections were rinsed with increasing stringency washes, subjected to RNase A (2 µg ml⁻¹, 30 min at 37 °C) and dehydrated before dipping in K₅ nuclear emulsion gel (Ilford). Autoradiographic exposure ranged from 21 to 28 days. The myogenin probe corresponded to nucleotides 31–638 of the coding sequence, whereas nucleotides 181–811 of the coding sequence were used for the M-cadherin probe. The myomaker probe was full-length coding sequence. ³⁵S-labelled sense and antisense probes were generated by Sp6 and T7 RNA polymerases, respectively, from linearized cDNA templates by *in vitro* transcription using a Maxiscript kit (Ambion).

Cardiotoxin injury. Cardiotoxin from *Naja mossambica mossambica* (Sigma) was dissolved in sterile saline to a final concentration of 10 µM, divided into aliquots and stored at -20 °C. Mice were anaesthetized by intraperitoneal injection of 2.5% Avertin at 15 µl g⁻¹. Mouse legs were shaved and cleaned with alcohol. Tibialis anterior muscles were injected with 50 µl of cardiotoxin with a 26-gauge needle.

X-gal staining. For whole-mount X-gal staining, either embryos or tissues were fixed in 4% PFA/PBS (containing 0.01% deoxycholic acid and 0.02% Igepal) for 45 min at 4 °C with gentle shaking then rinsed twice with cold PBS. Samples were stained overnight in staining solution (5 mM K₃Fe(CN)₆, 5 mM K₄Fe(CN)₆, 2 mM MgCl₂, 1 mg ml⁻¹ X-gal in PBS) followed by washing twice in PBS and post-fixing with 4% PFA/PBS.

For X-gal staining of cryosections or cells in culture, the following procedure was used: fix with 2% glutaraldehyde/PBS, wash three times in 0.1% sodium deoxycholate, 0.2% NP40 Substitute (Fluka), PBS, and incubate in staining solution

(4 mM K₃Fe(CN)₆, 4 mM K₄Fe(CN)₆, 0.4 mM MgCl₂, 1 mg ml⁻¹ X-gal, 0.1% sodium deoxycholate, 0.2% NP40 Substitute in PBS) at 37 °C overnight in the dark. The samples were then rinsed in PBS and fixed in 4% PFA/PBS for at least 20 min. Tissue sections were co-stained with light eosin, dehydrated and mounted with Permount (Fisher). Cells were co-stained with Nuclear Fast Red (Sigma).

Northern blot analysis. Total RNA was extracted as previously described. Fifteen micrograms of RNA were extracted, resolved on a 1% agarose/MOPS (0.2 M MOPS pH 7.0, 20 mM sodium acetate, 10 mM EDTA pH 8.0) gel and transferred to Hybond N+ membrane (Amersham). The membrane was then incubated in hybridization buffer (1% crystalline BSA (fraction V), 1 mM EDTA, 0.5 M NaHPO₄, 7% SDS) for at least 2 h at 68 °C followed by overnight incubation with probes labelled with [α -³²P]dCTP using the RadPrime DNA Labelling System (Invitrogen). Myomaker probe was generated from full-length coding sequence. The next day the membrane was washed with 1× SSC, 0.1% SDS for 10 min at room temperature followed by three washes at 68 °C with 0.5× SSC, 0.1% SDS. The membrane was exposed to film at -80 °C overnight and developed with a SRX101A Tabletop X-Ray Film Processor (Konica Minolta).

Histology and immunohistochemistry. For cryosections, skeletal muscle or limbs were dissected, embedded in gum tragacanth (1% in PBS), and frozen in 2-methylbutane-cooled liquid nitrogen. For paraffin sections, tissue was fixed in 10% neutral buffered formalin and processed for routine paraffin histology. Frozen and paraffin sections were cut and stained with H&E using routine procedures. Immunohistochemistry was performed by fixation with 1%PFA/PBS, permeabilization with 0.2% Triton X-100 in PBS, blocking with PBS/1% BSA, 1% heat inactivated goat serum, 0.025% Tween20, incubation with primary antibody for at least 2 h, incubation with secondary Alexa-Fluor antibodies (Invitrogen) for 1 h, and mounting with VectaShield containing DAPI (Vector Laboratories). Anti-mouse myosin (my32, Sigma) and desmin (DAKO) antibodies were used at 1:100. The TdT-mediated dUTP nick end labelling (TUNEL) (Invitrogen) reaction was performed exactly as described by the manufacturer. Slides were visualized using a Leica DM RXE microscope.

Isolation of primary myoblasts and immunocytochemistry. Limbs were dissected from E15 to E17.5 embryos and dissociated in 0.05% Collagenase D (Roche) in PBS at 37 °C for 2–3 h. Ten millilitres of culture media (20% FBS/Ham F10) were added to the suspension and triturated followed by centrifugation at 1,500g for 10 min at 4 °C. The pellet was re-suspended in 10 ml of growth media (20% FBS/Ham F10 + 2.5 ng ml⁻¹ bFGF (Promega)), filtered through a 100 µm cell strainer and plated on a 10 cm laminin-coated culture dish. To enrich for myoblasts, cultures were incubated in a small volume of PBS; the myoblasts were dislodged by knocking the plate lightly. To induce myogenesis, the cultures were placed in differentiation media (2% horse serum, DMEM) for 3–5 days. Immunocytochemistry was performed by fixing with 4% PFA/PBS, permeabilization with 0.2% Triton X-100 in PBS, blocking with 3% BSA/PBS, incubation with primary antibody for at least 2 h, then incubation with Alexa-Fluor secondary antibodies for 1 h. Myosin antibody, used as described above, M2 Flag antibody (Sigma) at 1:500, BrdU (Roche) at 1:100, EEA1 (a gift from the Schmid laboratory, University of Texas Southwestern Medical Center) at 1:500, GM130 (BD Pharmingen) at 1:300, cyclophilin D (Abcam) at 1:200, PDI (Cell Signaling) at 1:500. Cultures were co-stained with Phalloidin-rhodamine (Invitrogen) at 1:200 and nuclei were stained with Hoechst (Invitrogen). To stain live cells, we first washed the cells with PBS and incubated them in blocking buffer (3% BSA/PBS) for 15 min. Primary antibody incubation was then performed on ice, followed by fixation with 4% PFA/PBS and incubation with secondary antibody. These cultures were visualized on a Zeiss LSM 780 confocal microscope or a Nikon Eclipse Ti fluorescent microscope.

Cloning, generation of retroviruses and C2C12 infection. We cloned myomaker coding sequence from postnatal day 0 WT tongue cDNA using the following primers: myomaker-F, 5'-ATGGGGACAGTTGTAGCCAA-3'; myomaker-R, 5'-TCAGACACAAGTCGACGAGA-3'. Myomaker-Flag was generated by independently cloning the regions immediately upstream (5' PCR product) and downstream (3' PCR product) of the site of Flag insertion. These products were used as templates, and myomaker-F and myomaker-R as primers, in a standard PCR sewing reaction to generate full-length myomaker-Flag.

Retroviral plasmid DNA was generated by subcloning myomaker and Flag-tagged myomaker cDNA into the retroviral vector pBabe-X³¹. GFP and dsRed retrovirus have been described previously³². Ten micrograms of retroviral plasmid DNA were transfected with FuGENE 6 (Roche) into Platinum E cells (Cell Biolabs), which were plated on a 10 cm culture dish at a density of 3 × 10⁶ cells per dish, 24 h before transfection. Forty-eight hours after transfection, viral media was collected, filtered through a 0.45 µm cellulose syringe filter and mixed with polybrene (Sigma) at a final concentration of 6 µg ml⁻¹. C2C12 myoblasts (obtained from the American Type Culture Collection) were plated on 35 mm culture dishes at a density of 3 × 10⁵ cells per dish 24 h before infection with viral media. Eighteen hours after infection, virus was removed, cells were washed with PBS and replaced

with differentiation media. These cultures were assayed between 1 and 5 days of differentiation. The actin inhibitors cytochalasin D (Sigma) and lantrunculin B (Sigma) were used at concentrations of 0.3 μ M and 0.1 μ M, respectively.

Subcellular fractionation and western blot analysis. To fractionate C2C12 cells into cytosol and membrane fractions, we first washed a 10 cm dish with cold PBS and lysed the cells by dounce homogenization in hypotonic buffer (10 mM Tris pH 8.0, 1 mM EDTA). The homogenate was centrifuged at 500g for 5 min to pellet nuclei and cell debris. The supernatant was centrifuged at 100,000g for 20 min to pellet membrane structures. The supernatant from this step was the cytosol fraction and the membrane fraction was solubilized in an equal volume of hypotonic buffer + 1% *n*-dodecyl β -D-maltoside (DDM, Sigma) for further analyses by immunoblotting. For analysis of whole-cell extracts, DDM solubilization was used (20 mM HEPES, 150 mM NaCl, 2 mM EDTA, 10% glycerol, 1% DDM). For immunoblotting, equal protein amounts were separated on a 12% SDS-PAGE, transferred to a PVDF membrane (Millipore), blocked in 5% milk in TBS-tween and incubated with primary antibodies. The following antibodies were used: M2 Flag (Sigma, 1:1,000), Gapdh (Millipore, 1:10,000), VDAC (Santa Cruz, 1:1,000), α -tubulin (Sigma, 1:1,000), myosin (my32, Sigma, 1:1,000) and myogenin (Developmental Studies Hybridoma Bank, 1:1,000).

Cell mixing. WT myoblasts were mixed with either myomaker^{+/-} or myomaker^{-/-} myoblasts in equal ratios (approximately 1×10^5 cells per genotype), plated on a well of a laminin-coated 12-well plate and induced to differentiate the next day. 10T1/2 fibroblasts were infected with either GFP- and empty-retrovirus or GFP- and myomaker-retrovirus for 18 h. After infection, cells were washed several times, then trypsinized and mixed with C2C12 myoblasts at a ratio of 1:1 (1×10^5 of each cell type) and plated on one well of a six-well plate in differentiation media. GFP and myosin expression was analysed 4 days after differentiation. A similar protocol was performed to assess incorporation of BrdU-labelled fibroblasts into myotubes, with minor modifications. 10T1/2 fibroblasts were incubated with BrdU (Roche)

at a final concentration of 10 μ M for 18 h. They were then infected with either empty-retrovirus or myomaker-retrovirus and mixed with C2C12 myoblasts that had been infected with dsRed-retrovirus.

Time-lapse microscopy. In Supplementary Movie 1, C2C12 myoblasts were infected with GFP and myomaker retrovirus. For Supplementary Movie 2, C2C12 myoblasts were infected with dsRed retrovirus and fibroblasts were infected with GFP and myomaker retrovirus. GFP and dsRed was visualized using a Perkin Elmer Ultraview spinning disk confocal microscope with a chamber for control of temperature and CO₂. Images were captured every 15 min using Volocity 5.4.0 software. Images were analysed and movies assembled using ImageJ.

Quantification and statistics. Each histological analysis of embryonic skeletal muscle was performed on four samples per genotype. The differentiation index was calculated as the percentage of nuclei in myosin-positive cells. The fusion index was calculated as the percentage of nuclei contained in myosin-positive myotubes. Structures had to contain at least two nuclei to be considered a myotube. To quantify fusion between WT myoblasts and either myomaker^{+/-} or myomaker^{-/-} myoblasts, we calculated the percentage of LacZ⁺ myotubes containing at least three nuclei. To quantify fusion between fibroblasts and myoblasts we calculated the percentage of GFP⁺ myosin⁺ cells or the percentage of BrdU⁺ myotube nuclei. For each quantification, at least three independent experiments were performed in duplicate and at least six random fields were imaged per sample. Data are presented as mean \pm s.e.m. Differences between groups were tested for statistical significance using an unpaired two-tailed Student's *t*-test. *P* < 0.05 was considered significant.

32. Shelton, J. M., Lee, M. H., Richardson, J. A. & Patel, S. B. Microsomal triglyceride transfer protein expression during mouse development. *J. Lipid Res.* **41**, 532–537 (2000).

LRG1 promotes angiogenesis by modulating endothelial TGF- β signalling

Xiaomeng Wang¹, Sabu Abraham¹, Jenny A. G. McKenzie¹, Natasha Jeffs¹, Matthew Swire¹, Vineeta B. Tripathi¹, Ulrich F. O. Luhmann², Clemens A. K. Lange^{2,3,4}, Zhenhua Zhai⁵, Helen M. Arthur⁵, James W. B. Bainbridge^{2,3}, Stephen E. Moss^{1*} & John Greenwood^{1*}

Aberrant neovascularization contributes to diseases such as cancer, blindness and atherosclerosis, and is the consequence of inappropriate angiogenic signalling. Although many regulators of pathogenic angiogenesis have been identified, our understanding of this process is incomplete. Here we explore the transcriptome of retinal microvessels isolated from mouse models of retinal disease that exhibit vascular pathology, and uncover an upregulated gene, leucine-rich alpha-2-glycoprotein 1 (*Lrg1*), of previously unknown function. We show that in the presence of transforming growth factor- β 1 (TGF- β 1), LRG1 is mitogenic to endothelial cells and promotes angiogenesis. Mice lacking *Lrg1* develop a mild retinal vascular phenotype but exhibit a significant reduction in pathological ocular angiogenesis. LRG1 binds directly to the TGF- β accessory receptor endoglin, which, in the presence of TGF- β 1, results in promotion of the pro-angiogenic Smad1/5/8 signalling pathway. LRG1 antibody blockade inhibits this switch and attenuates angiogenesis. These studies reveal a new regulator of angiogenesis that mediates its effect by modulating TGF- β signalling.

The formation of new blood vessels by angiogenesis is a key feature of several diseases including age-related macular degeneration, proliferative diabetic retinopathy (PDR), atherosclerosis, rheumatoid arthritis and cancer. The factors that promote neovascularization have been the subject of extensive research, with the vascular endothelial growth factors (VEGFs) and their receptors emerging as master regulators^{1–3}. Despite the prominent role of VEGF, other factors contribute to neoangiogenesis through coordinated crosstalk that is often highly context-dependent^{4–6}. Such complexity is exemplified in TGF- β 1 signalling, which can switch from being mostly angiostatic to pro-angiogenic⁷. What regulates this switch is not fully understood, but activation of the pro-angiogenic pathway involves TGF- β type II receptor (T β RII) recruitment of the predominantly endothelial TGF- β type I receptor activin receptor-like kinase-1 (ALK1), which in turn initiates activation of the transcription factors Smad1, 5 and 8, resulting in a pro-angiogenic phenotype^{7–10}. The regulation of this differential signalling is contingent on several factors including the concentration of TGF- β , its bioavailability and the presence or absence of other regulatory factors such as bone morphogenic proteins and accessory receptors such as endoglin (ENG) and betaglycan (also known as TGF- β type III receptor)¹¹.

Our incomplete understanding of the role of the fine-tuning of angiogenesis suggests that additional modulators have yet to be identified. Our objective in this study, therefore, was to identify new regulators of pathogenic angiogenesis that may lead to the development of more effective treatment strategies.

Retinal vascular expression of LRG1

To identify new regulators of neovascularization we exploited three mouse mutants that exhibit marked remodelling of the retinal vasculature (Supplementary Fig. 1 and Supplementary Videos 1–4). Genome-wide transcriptome analysis of retinal microvessel fragments isolated from the retinal degeneration 1 (*rd1*) mouse, the very low density lipoprotein receptor

(VLDLR) knockout mouse (*Vldlr*^{−/−}), the *Grhl3*^{ct/J} curly tail mouse (Jackson Laboratory) and appropriate wild-type control mice yielded 62 genes that were differentially regulated but common to all three retinal disease models (Supplementary Table 1). When ranked according to fold change, a gene encoding a secreted glycoprotein of unknown function, namely *Lrg1*, emerged as the most significantly upregulated. LRG1 is a highly conserved member of the leucine-rich repeat family of proteins, many of which are involved in protein–protein interactions, signalling and cell adhesion (Supplementary Fig. 2a, b).

Validation of the microarray data revealed that in the retina, LRG1 is restricted almost exclusively to the vasculature, is expressed under normal conditions and is upregulated during retinal vascular remodelling in the three mouse models of retinal disease (Fig. 1a–d and Supplementary Fig. 3). However, LRG1 expression was not restricted to the retina, as we also observed LRG1 staining in the choriocapillaris of the mouse eye (Supplementary Fig. 4a). Consistent with the data obtained in the mouse, we observed low levels of constitutive LRG1 expression in normal adult human retinal vessels and weakly, but not exclusively, in vessels in other human tissues including breast, skin and intestine (Supplementary Fig. 4b).

We next investigated whether the *Lrg1* transcript is also increased in the retinæ of models of choroidal and retinal neovascularization. Choroidal neovascularization (CNV) was induced in wild-type mice, and 1 week after laser injury we observed a significant increase in *Lrg1* transcript levels in both the retina and retinal pigment epithelium (RPE)/choroid (Fig. 1e, f). We then examined intra-retinal/pre-retinal neovascularization in the mouse model of oxygen-induced retinopathy (OIR), which displays hypoxia-driven retinal angiogenesis. At postnatal day (P) 17, during the ischaemic proliferative phase of OIR when neovascularization is most prevalent, *Lrg1* transcript levels were also upregulated (Fig. 1g). However, at the end of the hyperoxic phase (P12), *Lrg1* messenger RNA was significantly reduced. Indeed, the pattern of *Lrg1* expression at the two time points observed mirrored

¹Department of Cell Biology, UCL Institute of Ophthalmology, London EC1V 9EL, UK. ²Department of Genetics, UCL Institute of Ophthalmology, London EC1V 9EL, UK. ³NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital, London EC1V 2PD, UK. ⁴University Eye Hospital Freiburg, Freiburg 79106, Germany. ⁵Institute of Genetic Medicine, Newcastle University, Newcastle NE1 3BZ, UK. *These authors contributed equally to this work.

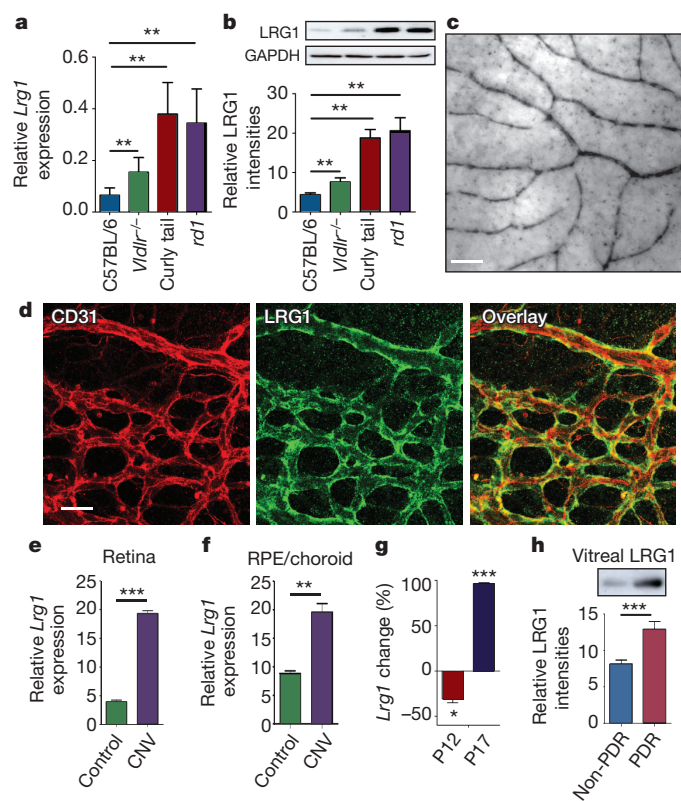


Figure 1 | LRG1 is overexpressed in pathogenic retinal vasculature.

a, b, Quantification of *Lrg1* mRNA (**a**) and LRG1 protein expression (**b**), showing upregulation in the retinas of mice exhibiting retinal vascular changes. **c**, *Lrg1* *in situ* hybridization at P21. Scale bar, 50 μ m. **d**, Immunohistochemical detection of CD31 (red) and LRG1 (green) at P10, showing LRG1 expression in the retinal vasculature. **e, f**, Upregulation of *Lrg1* mRNA in the retina (**e**) and RPE/choroid (**f**) in CNV mice. **g**, Reduced *Lrg1* transcript levels in OIR at P12 and increased levels at P17. **h**, Increase in LRG1 protein in the vitreous of patients with PDR. All images shown are representative and data are mean \pm s.e.m. of $n \geq 3$ independent experimental groups. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test).

the expression of the hypoxia-responsive genes *Vegfa*, *Apln* (apelin) and its receptor *Aplnr* (Supplementary Fig. 5). To determine whether LRG1 is upregulated in human retinal disease in which there is neo-vascular pathology, vitreous samples from human subjects with PDR were analysed by western blot, which revealed increased LRG1 expression compared to control vitreous (Fig. 1h and Supplementary Fig. 6). It is unclear, however, whether this increase is the consequence of increased local production, leakage from the systemic circulation or a combination of both.

These data show that in the retina, LRG1 expression is predominantly vascular, is constitutive, and is increased during neovascular growth.

LRG1 and angiogenesis

To investigate the function of LRG1 we used cultured endothelial cell assays and *in vitro* and *ex vivo* models of angiogenesis. We observed that overexpression of human LRG1 in endothelial cells increased proliferation, whereas knockdown of mouse *Lrg1* decreased proliferation (Supplementary Fig. 7). In addition, endothelial cell migration was inhibited by an anti-LRG1 polyclonal antibody (Supplementary Figs 7 and 8). In the Matrigel human umbilical vein endothelial cell (HUVEC) tube-formation assay, the supplementation of media with recombinant human LRG1 (Supplementary Fig. 8) caused a significant increase in tube formation and branching, whereas an anti-LRG1 antibody significantly blocked tube formation (Fig. 2a and Supplementary Figs 8 and 9). Consistent with the latter observation, LRG1 was

found to be present in the conditioned media of these assays (Supplementary Fig. 10). We next investigated whether LRG1 promotes blood vessel growth in two *ex vivo* models of angiogenesis. Mouse metatarsals (embryonic day (E) 16.5) and aortic rings (P7) were prepared using tissues from wild-type mice. Vessel outgrowth and branching from explanted metatarsals (Supplementary Fig. 11) or aortic rings in the absence of other added growth factors were significantly increased after the addition of exogenous LRG1, and inhibited in the presence of the anti-LRG1 polyclonal antibody (Fig. 2b). Again, conditioned media from both assays was found to contain LRG1 protein (Supplementary Fig. 10).

Having demonstrated that LRG1 influences vascular growth *in vitro* and *ex vivo* we then investigated the retinal vasculature of the *Lrg1* knockout mouse (Supplementary Fig. 12). *Lrg1*^{-/-} mice were viable but exhibited a delay in the development of the deep vascular plexus at P10–P12 and the intermediate vessels between P17 and P25 that had resolved by P35 (Supplementary Fig. 13). In addition, the hyaloid vessels failed to regress fully, with vessel persistence beyond

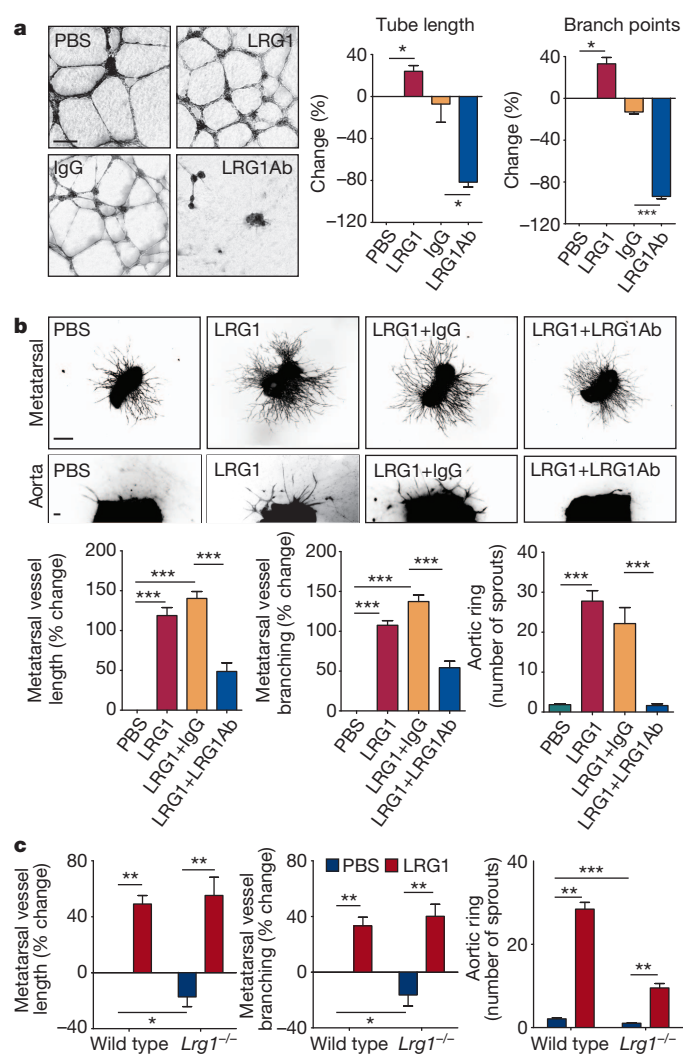


Figure 2 | LRG1 promotes angiogenesis. **a**, Increased HUVEC tube and branch formation after the addition of LRG1, and inhibition by a LRG1 neutralizing antibody (LRG1Ab). Scale bar, 160 μ m. **b**, Vessel outgrowth in the metatarsal (top) and aortic ring (bottom) assay is enhanced by LRG1 and attenuated by a LRG1 neutralizing antibody. Scale bar, 1,500 μ m. **c**, Comparison of vessel growth from metatarsals and aortic rings isolated from wild-type and *Lrg1*^{-/-} mice shows reduced angiogenesis in the latter that could be rescued by the addition of LRG1. All images shown are representative and values are expressed as mean \pm s.e.m. of $n \geq 3$ independent experimental groups. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test).

P35 and integration into the inner retina (Supplementary Fig. 14). Defective retinal vascular development and persistent hyaloid vessels were also reported in mice with deletions in *Ndp* (Norrie disease (pseudoglioma), also known as norrin), *Fzd4* (frizzled homolog 4), *Lrp5* and *Angpt2* (angiopoietin 2), which also contribute to angiogenesis^{12–14}. We also observed an increase in the incidence of crossover of the radial arteries and veins and of their side branches, with occasional small vessels forming arteriovenous anastomosis (Supplementary Fig. 15). Arteriovenous crossing has been reported in the retina of the hypomorphic *Vegfa* mouse¹⁵ and is associated with susceptibility to branched vein occlusion in the human retina^{16,17}. In this context it was interesting to note that *Vegfa* gene expression in the *Lrg1*^{−/−} mouse retina is significantly lower than in control mice in contrast to *Plgf*, which is unchanged (Supplementary Fig. 16). Aside from these mild defects, the retinal vasculature of the *Lrg1*^{−/−} mice exhibited similar pericyte

coverage (Supplementary Fig. 17) and barrier properties (Supplementary Fig. 18) to wild-type controls.

As we had observed that LRG1 inhibition or supplementation had a significant effect on vessel formation in the metatarsal and aortic ring assays, we proposed that *Lrg1* knockout would lead to reduced angiogenesis in these models. Indeed, vessel formation was significantly reduced in *Lrg1*^{−/−} mice in both the metatarsal and aortic ring assay (Fig. 2c and Supplementary Fig. 19), and could be rescued by the addition of exogenous LRG1. Together, these data support the hypothesis that LRG1 contributes to, and is necessary for, robust vascular growth.

LRG1 and pathogenic neovascularization

As our data thus far had demonstrated increased *Lrg1* transcript expression in CNV and OIR in wild-type mice, we investigated whether neovascularization in these models is attenuated in *Lrg1*^{−/−} mice. CNV was induced in wild-type and *Lrg1*^{−/−} mice, and at 7 days post-laser fundus fluorescein angiography at 90 s revealed a diminished neovascular response in the *Lrg1*^{−/−} mice compared to controls (Fig. 3a; $P < 0.01$). Concomitant with this was an equivalent reduction in fluorescein leakage at 7 min after injection. This effect was confirmed in a group of animals in which the neovascular lesion was visualized in posterior eyecup whole mounts, quantitative analysis of which showed that mean lesion volume was about 70% smaller in *Lrg1*^{−/−} than wild-type mice (Supplementary Fig. 20). The reduction in lesion size in the *Lrg1*^{−/−} mouse was similar to that reported in *Plgf*^{−/−} (ref. 18) and *Ccr3*^{−/−} (ref. 19) mice, two other pro-angiogenic factors, and could not be explained by changes in macrophage recruitment (Supplementary Fig. 20) or pericyte coverage (Supplementary Fig. 21). We next investigated intra-retinal/pre-retinal neovascularization, as observed in PDR, in the OIR model of angiogenesis. After the 5-day hyperoxia phase, the size of the avascular region at P12 was not significantly different between the *Lrg1*^{−/−} and wild-type animals (Fig. 3b and Supplementary Fig. 22). Furthermore, after 5 days in normoxia, revascularization of the avascular region with ordered vessels was similar between the two groups, demonstrating that hyperoxia-induced regression and hypoxia-induced physiological revascularization are not affected by the loss of LRG1. However, the area occupied by disordered neovascular growth (tufts) was significantly reduced in the absence of LRG1 (Fig. 3b; $P < 0.01$; Supplementary Fig. 22), demonstrating that LRG1 is specifically required for robust pathogenic angiogenesis.

Having demonstrated that an antibody against LRG1 inhibits angiogenesis *in vitro*, we investigated whether this antibody would also reduce CNV lesion size. After the laser burn, animals received intravitreal injections of the anti-LRG1 polyclonal antibody or a pre-immune IgG as a control, and 5 days later lesion sizes were measured. In the anti-LRG1 antibody-treated eyes, a dose-dependent reduction in CNV lesion volume (Fig. 3c, d) and area (Supplementary Fig. 23) was observed compared to control-antibody-treated eyes. Indeed, the 58% reduction of CNV volume (Fig. 3d) and 46% reduction in area at a dose of 10 μ g (Supplementary Fig. 23) was of similar magnitude as that achieved with blockade of the VEGF/PLGF signalling axis²⁰ or the chemokine receptor CCR3 (ref. 19).

As antibody blockade of LRG1 reduces CNV lesion size, we investigated the effects of combination therapy with a VEGF receptor 2 (VEGFR2)-blocking antibody²⁰. Antibody blockade of LRG1, VEGFR2, or both LRG1 and VEGFR2 together inhibited CNV lesion volume (Fig. 3e and Supplementary Fig. 23), with the combined therapy giving the most significant inhibition. The effect of combinatorial treatment was also evaluated in OIR. Under the treatment conditions used, in which the individual antibodies elicited no significant effect, animals treated with the antibody combination exhibited a significant inhibition of both patterned revascularization and the formation of pathogenic vascular tufts (Supplementary Fig. 24). These data provide compelling evidence that inhibition of LRG1 is effective in preventing pathological

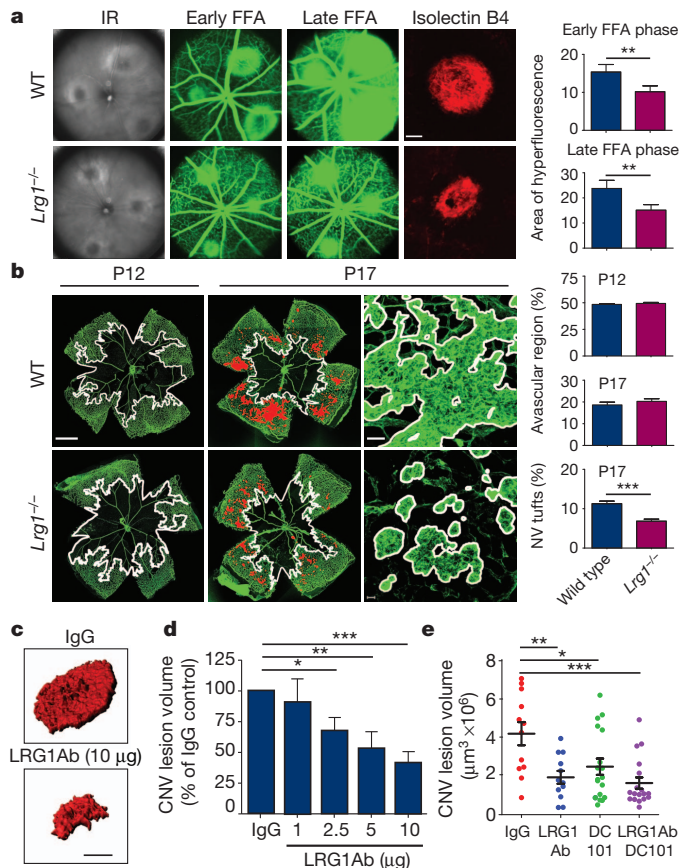


Figure 3 | LRG1 contributes to pathogenic neovascularization.

a, Representative images of wild-type (WT) and *Lrg1*^{−/−} mouse laser-burn lesions by infrared (IR) fundus imaging. At 7 days after the laser, early- and late-phase fundus fluorescein angiography (FFA) revealed a reduction in CNV lesion size and a decrease in fluorescein leakage, respectively, in *Lrg1*^{−/−} mice. Representative images of isolectin B4 stained (red) CNV in choroidal/RPE flat-mount 7 days after induction, confirming decreased lesion size in *Lrg1*^{−/−} mice. Scale bar, 100 μ m. **b**, In OIR, *Lrg1* deletion does not affect the size of the avascular region at P12 (delineated by white boundary line) or the organized normal revascularization at P17, but does decrease the formation of pathological neovascular (NV) tufts (highlighted in red and delineated in higher power by white boundary line). Scale bars, 1,000 μ m (P12 and P17 lower magnification) and 50 μ m (P17 higher magnification). **c**, Volume-rendered examples of PECAM-1 stained CNV lesions in wild-type mice after intravitreal injection of irrelevant IgG or LRG1 neutralizing antibody. Scale bar, 100 μ m. **d**, Dose-dependent anti-LRG1 antibody reduction of CNV lesion volume. **e**, Combination of anti-LRG1 and DC101 (anti-VEGFR2) in CNV in wild-type mice resulted in enhanced reduction of lesion volume compared to single treatments. Data are mean \pm s.e.m. of $n \geq 10$ for each group. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test (a, b), and one-way analysis of variance (ANOVA)(d, e)).

angiogenesis, and suggest that LRG1 has potential as a therapeutic target on its own or in combination with other anti-angiogenic therapies.

LRG1 and TGF- β signalling

Although little is known about the biology of LRG1, concomitant increases in the expression levels of TGF- β 1, T β RII and LRG1 have been reported in cancer cells²¹ and hydrocephalus²², and LRG1 has been shown to bind to TGF- β 1 in high endothelial venules²³. Consistent with this, and with earlier proteomic and transcriptome analyses^{24,25}, we have shown here that in vitreous samples from the eyes of human subjects with PDR, both LRG1 (Fig. 1h and Supplementary Fig. 6) and TGF- β 1 protein levels (Supplementary Fig. 25) are significantly increased. Furthermore, alongside increased *Lrg1* gene expression, *Tgfb1* transcript levels were also significantly upregulated in the retinae of laser-induced CNV mice and in OIR mice during the ischaemic proliferative phase (Supplementary Fig. 5). These data prompted us to investigate whether LRG1 acts as a modulator of TGF- β signalling.

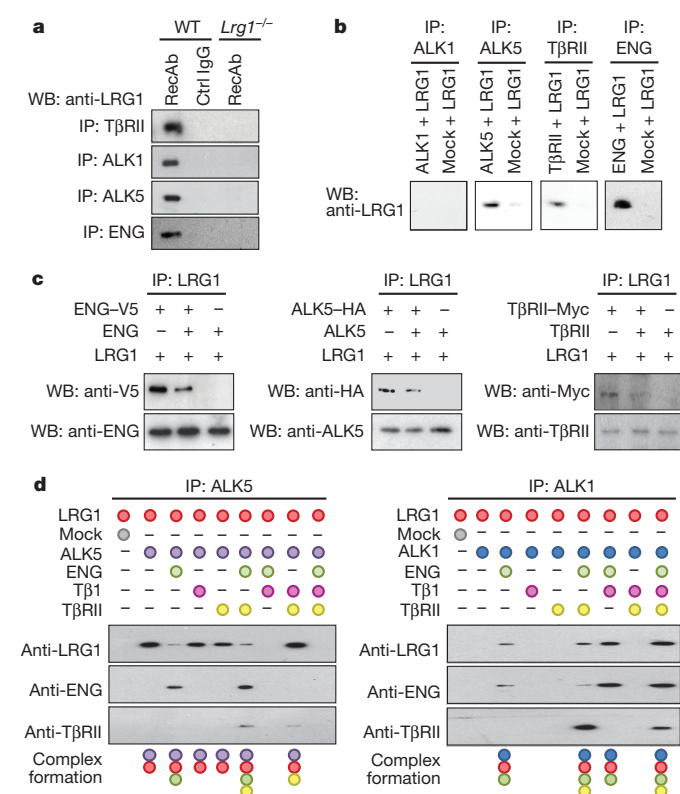


Figure 4 | LRG1 modifies the TGF- β receptor complex.

a, Immunoprecipitation (IP) of T β RII, ALK1, ALK5 and ENG with anti-receptor antibodies (RecAb) from wild-type mouse brain endothelial cell lysates co-precipitates LRG1. Control IgG in wild-type endothelial cells or anti-receptor antibodies in *Lrg1*^{-/-} endothelial cells did not co-precipitate LRG1. WB, western blot. **b**, Immunoprecipitation of peptide-tagged extracellular domains of ALK5 (haemagglutinin (HA)-tagged), T β RII (Myc-tagged) or ENG (V5-tagged) added individually to histidine (His)-tagged LRG1 resulted in co-precipitation of LRG1, indicating direct interactions with these receptors. Immunoprecipitation of ALK1 (HA-tagged) in the presence of LRG1 did not co-precipitate the latter. **c**, Addition of appropriate soluble non-tagged extracellular domains of ENG, ALK5 and T β RII out-competed peptide-tagged receptor binding to LRG1. **d**, LRG1 was incubated *in vitro* with different combinations of TGF- β receptor extracellular domains and TGF- β 1 (T β 1). In the presence of ENG, binding between LRG1 and ALK5 is diminished, and is completely lost with the further addition of TGF- β 1. Conversely, ENG facilitates the association between LRG1 and ALK1, which is enhanced in the presence of TGF- β 1. Although T β RII has no effect on LRG1–ALK1 or LRG1–ALK5 interactions, it is recruited to the complex in the presence of ENG. All data are representative western blots of $n \geq 3$ for each experiment.

To determine whether LRG1 associates with components of the TGF- β receptor complex, we performed co-immunoprecipitation experiments. In primary brain endothelial cells, LRG1 was present in immunoprecipitates of T β RII, ALK1, ALK5 and the auxiliary receptor ENG (Fig. 4a). Conversely, immunoprecipitates of LRG1 from HUVECs were found to contain T β RII, ALK1, ALK5 and ENG (Supplementary Fig. 26a). These observations suggested that LRG1 might be involved in regulating TGF- β signalling through fine-tuning the stoichiometry of the TGF- β receptor complex. TGF- β 1 signals in endothelial cells by T β RII recruitment of either the ubiquitous ALK5 receptor, or the predominantly endothelial ALK1 receptor together with ALK5. Stimulation of the T β RII–ALK5 signalling complex results in phosphorylation and activation of the transcription factors Smad2 and 3, which increases extracellular matrix deposition, inhibits endothelial cell proliferation and migration, and promotes cell homeostasis, whereas signalling via the T β RII–ALK5/ALK1 complex (possibly in association with ENG) activates Smad1, 5 and 8, resulting in a pro-angiogenic state^{7,8}. We therefore investigated the direct one-to-one binding of LRG1 to recombinant extracellular domains of individual TGF- β receptors in serum-free conditioned media from transfected HEK293T cells (Supplementary Fig. 26b). Immunoprecipitation of the receptor ectodomain revealed co-immunoprecipitation of LRG1 with ALK5, T β RII and ENG, indicating a direct interaction of LRG1 with these individual receptors (Fig. 4b). This occurred in the absence of TGF- β 1, which was not present in HEK293T-cell-conditioned medium (Supplementary Fig. 10). Addition of conditioned medium containing non-tagged ENG, ALK5 or T β RII out-competed tagged receptor binding to LRG1 (Fig. 4c), confirming the specificity of these protein–protein interactions. The observation that ENG seems to be one of the receptors for LRG1 is germane, given its proposed role in switching TGF- β signalling towards the pro-angiogenic Smad1/5/8 pathway²⁶. A potential functional relationship between LRG1 and ENG was additionally strengthened by our observation that *Eng* is upregulated in CNV and OIR (Supplementary Fig. 27). To define the LRG1–ENG interaction further we undertook surface plasmon resonance analysis (Biacore) and obtained an affinity rate constant (K_D) of 2.9 μ M (Supplementary Fig. 28) for binding of the ENG ectodomain to LRG1.

These data raise the possibility that LRG1 facilitates a receptor configuration conducive to the pro-angiogenic signalling pathway. To investigate this, LRG1 was incubated with conditioned media containing the extracellular domains of either ALK1 or ALK5 in the presence or absence of different combinations of TGF- β 1, and the extracellular domains of ENG and T β RII. These studies revealed that LRG1 only associated with ALK1 in the presence of ENG, to which it bound, and this was enhanced by the addition of TGF- β 1 (Fig. 4d). Conversely, ALK5 binds LRG1 in the absence of ENG but in its presence this interaction is attenuated, suggesting competition between ENG and ALK5 for LRG1, whereas the addition of TGF- β 1 results in complete loss of the LRG1–ALK5 association. Moreover, neither TGF- β 1 nor T β RII on their own, nor a combination of both, affects ALK1–LRG1 or ALK5–LRG1 association in the absence of ENG. However, in the presence of ENG, T β RII is able to form a complex with ALK1–LRG1 or ALK5–LRG1, with the former association being enhanced and the latter being further inhibited by TGF- β 1. In accordance with previously suggested models⁸, these data indicate that LRG1 may be able to form an intermediate complex with ALK5, ALK1, T β RII and ENG, but that in the presence of TGF- β 1 the LRG1–ALK1–T β RII–ENG complex predominates. An association between LRG1 and TGF- β 1 may therefore lead to more efficient ALK1–T β RII–ENG receptor complex formation, and consequently to the promotion of pro-angiogenic Smad1/5 signalling.

To test this hypothesis, we treated mouse brain endothelial cells with TGF- β 1 (5 ng ml⁻¹) and showed that both Smad2/3 and Smad1/5 phosphorylation are induced in wild-type cells, but in *Lrg1*-null cells only Smad2/3 is activated (Fig. 5a). The addition of LRG1 alone did

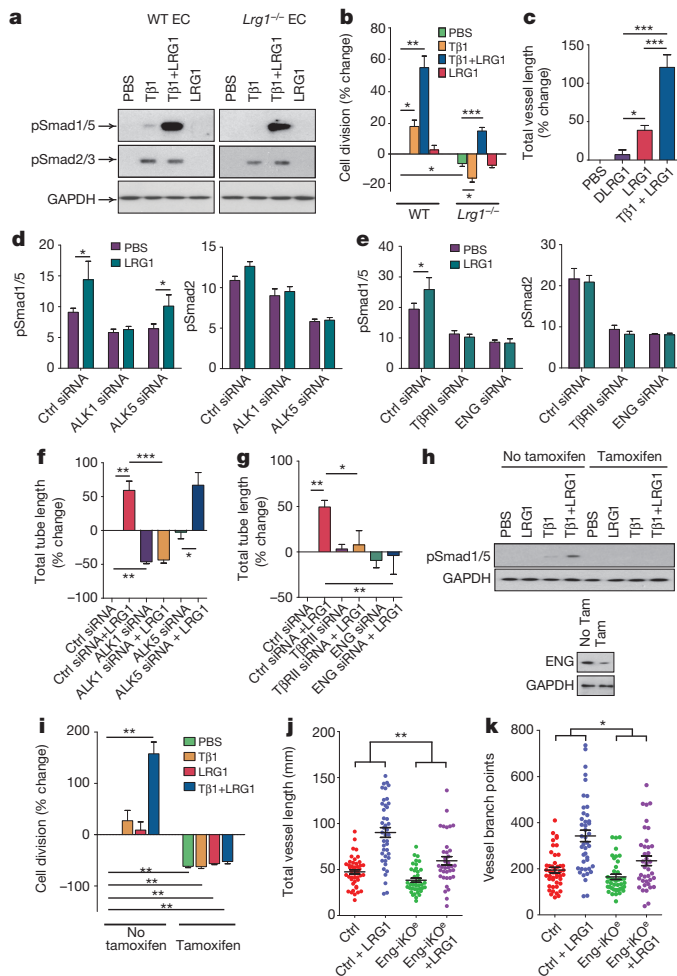


Figure 5 | LRG1 promotes angiogenesis via a switch in TGF- β signalling.

a, In wild-type brain endothelial cells (EC), TGF- β 1 stimulates Smad2/3 phosphorylation (pSmad2/3) and low levels of Smad1/5 phosphorylation (pSmad1/5), but in *Lrg1*^{-/-} cells only Smad2/3 is phosphorylated. LRG1 addition has no effect on Smad phosphorylation in wild-type or *Lrg1*-null cells, but co-treatment with TGF- β 1 and LRG1 enhances Smad1/5 phosphorylation without affecting Smad2/3 phosphorylation ($n \geq 3$). Ctrl, control. **b**, Proliferation of brain endothelial cells isolated from wild-type control and *Lrg1*^{-/-} mice after exogenous TGF- β 1 and/or LRG1 treatment normalized to control ($n \geq 3$). Non-treated *Lrg1*^{-/-} cells are less proliferative than wild-type cells. TGF- β 1 addition to wild-type cells results in enhanced proliferation but reduces proliferation in *Lrg1*^{-/-} cells, whereas TGF- β 1 and LRG1 co-treatment results in enhanced proliferation in wild-type and *Lrg1*^{-/-} cells. **c**, Addition of exogenous TGF- β 1 and LRG1, compared to LRG1 alone or denatured (D)LRG1, enhances microvessel formation in the mouse metatarsal angiogenesis assay ($n = 3$ independent experiments, $n \geq 30$ metatarsals per treatment). **d**, siRNA knockdown of ALK1 or ALK5 in HUVECs results in reduced Smad1/5 or Smad2 phosphorylation, respectively. ALK1, but not ALK5, knockdown results in prevention of LRG1-induced Smad1/5 phosphorylation. **e**, siRNA knockdown of T β RII or ENG inhibits LRG1-induced Smad1/5 phosphorylation. Histograms in **d** and **e** show semi-quantification of Smad phosphorylation relative to GAPDH ($n \geq 3$). **f**, **g**, Knockdown of ALK1, T β RII or ENG, but not ALK5, reduces LRG1-mediated HUVEC Matrigel tube formation ($n = 3$ independent groups for each assay). **h**, Treatment of lung endothelial cells isolated from *Rosa26-CreERT:Eng*^{+/+} mice (MLEC;Eng^{+/+}) with a combination of TGF- β 1 and LRG1 results in Smad1/5 phosphorylation; this response is lost after pre-treatment with 4OH-tamoxifen to delete ENG (MLEC;Eng^{-/-}). **i**, Treatment of control MLEC;Eng^{+/+} with TGF- β 1 and LRG1 stimulates cell division. In MLEC;Eng^{-/-} cells, cell division is reduced and refractive to treatment with TGF- β 1 \pm LRG1 ($n = 3$ independent experiments). **j**, **k**, 4OH-tamoxifen treatment of metatarsals isolated from *Eng*^{+/+} (control) and *Cdh5(PAC)-CreERT2:Eng*^{+/+} (Eng-iKO⁺) mice results in loss of ENG expression in the latter (Supplementary Fig. 30), and decreases LRG1-induced metatarsal vessel length (**j**) and branching (**k**) (metatarsals from five independent litters). Data are mean \pm s.e.m. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ (Student's *t*-test (**a**–**i**) and two-way ANOVA (**j** and **k**)).

not activate Smad2/3 or Smad1/5, but in combination with TGF- β 1 there was a marked induction of Smad1/5 phosphorylation, showing that LRG1 requires the presence of TGF- β 1 to stimulate the pro-angiogenic T β RII–ALK1–Smad1/5/8 pathway (Fig. 5a). As TGF- β 1 is an endothelial cell mitogen⁷ we also investigated whether LRG1 augmented TGF- β 1-mediated cell proliferation. Brain endothelial cells from *Lrg1*^{-/-} mice proliferated more slowly than those from wild-type animals (Fig. 5b). Addition of TGF- β 1 significantly enhanced endothelial cell proliferation from wild-type animals but inhibited the growth of cells from *Lrg1*^{-/-} mice, presumably through enhanced ALK5–Smad2/3 signalling in the absence of activation of the ALK1–Smad1/5/8 arm. The addition of LRG1 on its own had no effect, but TGF- β 1 and LRG1 in combination increased proliferation significantly in both wild-type and *Lrg1*-null endothelial cells (Fig. 5b). Moreover, in the metatarsal angiogenesis assay, the combined addition of LRG1 and TGF- β 1 led to a substantial increase in vessel formation (Fig. 5c). The observation that LRG1 alone induced a small increase in vessel formation indicates that TGF- β 1 is produced constitutively by the metatarsal tissue, which was confirmed by western blotting of conditioned medium (Supplementary Fig. 10).

To confirm that the pro-angiogenic effect of LRG1 was mediated through the ALK1–Smad1/5/8 pathway, ALK1 was knocked down with short interfering RNA (siRNA) (Fig. 5d and Supplementary Fig. 29) or inhibited by LDN193189 (Supplementary Fig. 29), resulting in prevention of LRG1-induced Smad1/5 phosphorylation without affecting Smad2 phosphorylation. As predicted, in the Matrigel assay ALK1 inhibition led to a significant decrease in HUVEC tube and branch formation, and blocked the angiogenic activity of LRG1 in this assay (Fig. 5f and Supplementary Fig. 29). Conversely, knockdown of ALK5 with siRNA (Fig. 5d, f) or inhibition with SB43152

(Supplementary Fig. 29), which inhibited constitutive Smad2 phosphorylation, did not prevent LRG1-induced HUVEC tube formation. siRNA knockdown of T β RII or ENG also resulted in the abrogation of LRG1-induced Smad1/5 phosphorylation (Fig. 5e and Supplementary Fig. 29) and HUVEC tube and branch formation (Fig. 5g). To corroborate the involvement of ENG in LRG1-mediated signalling further, lung endothelial cells derived from *Rosa26-CreERT:Eng*^{+/+} mice²⁷ were treated with 4OH-tamoxifen to deplete ENG (MLEC;Eng^{-/-}, in which MLECs denotes mouse lung endothelial cells) (Fig. 5h). Unlike control cells, treatment of MLEC;Eng^{-/-} with a combination of TGF- β 1 and LRG1 failed to induce Smad1/5 phosphorylation (Fig. 5h). Consistent with this, similar treatment resulted in a significant increase in cell division of control MLECs, whereas the cell division of MLEC;Eng^{-/-} was significantly reduced and refractive to treatment with TGF- β 1 \pm LRG1 (Fig. 5i). In addition, we used an endothelial-specific conditional knockout approach in which metatarsals were collected from *Eng*-floxed mice (*Cdh5(PAC)-CreERT2:Eng*^{+/+})²⁸ and treated with 4OH-tamoxifen (to generate Eng-iKO⁺ metatarsals). This resulted in a loss of vascular ENG expression, compared with *Eng*^{+/+} controls (Supplementary Fig. 30), and a 51% reduction in LRG1-induced vessel growth (Fig. 5j; $P < 0.01$) and a 53% reduction in vessel branching (Fig. 5k; $P < 0.05$). In agreement with these data, there were fewer cells expressing phosphorylated Smad1/5/8 in CNV lesions in *Lrg1*^{-/-} mice compared to wild-type animals (Supplementary Fig. 31). Moreover, in both CNV and OIR (during the neovascular phase) the Smad1/5 mediated pro-mitogenic gene inhibitor of DNA binding 1 (*Id1*) was significantly up-regulated (Supplementary Fig. 5).

Conclusions

TGF- β signalling has an important role in determining endothelial cell function during both development and vascular pathology^{7,8,29,30},

and its activity is regulated at several levels from gene expression to control of extracellular bioavailability. The multiplicity of regulatory mechanisms together with variable combinations of receptors/co-receptors creates complex patterns of TGF- β activity that define its context-dependent effects. In particular, the balance between the ALK5 and ALK1 signalling pathways is considered to be central in determining the angiogenic switch, with ENG being proposed as a key regulatory molecule in promoting signalling through the ALK1 pathway^{26,30,31}. In searching for mediators of vascular remodelling in the diseased/damaged retina we have discovered a new regulator of TGF- β signalling. The data presented here support a hypothesis that LRG1 activates the TGF- β angiogenic switch by binding to the accessory receptor ENG and, in the presence of TGF- β 1, promotes signalling via the T β RRII–ALK1–Smad1/5/8 pathway (Supplementary Fig. 32). Moreover, our evidence suggests that LRG1 may have a more dominant role in disorganized pathological rather than developmental/physiological angiogenesis. Although in the retina this is clearly supported by our *in vivo* data, the *ex vivo* and *in vitro* studies indicate that LRG1 angiogenic activity is not restricted to the eye. The modulating effect of LRG1 on TGF- β 1 signalling is the first demonstration, to our knowledge, of a definitive function for LRG1 and raises the intriguing possibility that it may influence other major biological processes in which TGF- β has a role, such as neoplasia³² and the immune response³³. Inhibition of LRG1, which we show here causes a shift away from angiogenic signalling, could prevent pathogenic activation of this pathway, while leaving homeostatic TGF- β signalling unperturbed. From these studies we suggest, therefore, that LRG1 is a highly promising therapeutic target for controlling pathogenic angiogenesis in ocular disease, and potentially in other diseases such as cancer and atherosclerosis.

METHODS SUMMARY

Microvessel global gene expression analysis was undertaken using Affymetrix mouse 430.2 gene arrays. CNV and OIR were induced in mice as described in the Methods. All other methods are described in the Methods.

Full Methods and any associated references are available in the online version of the paper.

Received 26 August 2011; accepted 3 June 2013.

- Leung, D. W., Cachianes, G., Kuang, W. J., Goeddel, D. V. & Ferrara, N. Vascular endothelial growth factor is a secreted angiogenic mitogen. *Science* **246**, 1306–1309 (1989).
- Carmeliet, P. *et al.* Abnormal blood vessel development and lethality in embryos lacking a single VEGF allele. *Nature* **380**, 435–439 (1996).
- Ferrara, N. *et al.* Heterozygous embryonic lethality induced by targeted inactivation of the VEGF gene. *Nature* **380**, 439–442 (1996).
- Holderfield, M. T. & Hughes, C. C. Crosstalk between vascular endothelial growth factor, notch, and transforming growth factor- β in vascular morphogenesis. *Circ. Res.* **102**, 637–652 (2008).
- Chung, A. S. & Ferrara, N. Developmental and pathological angiogenesis. *Annu. Rev. Cell Dev. Biol.* **27**, 563–584 (2011).
- Carmeliet, P. & Jain, R. K. Molecular mechanisms and clinical applications of angiogenesis. *Nature* **473**, 298–307 (2011).
- Pardali, E., Goumans, M. J. & ten Dijke, P. Signaling by members of the TGF- β family in vascular morphogenesis and disease. *Trends Cell Biol.* **20**, 556–567 (2010).
- Goumans, M. J., Liu, Z. & ten Dijke, P. TGF- β signaling in vascular biology and dysfunction. *Cell Res.* **19**, 116–127 (2009).
- Cunha, S. I. *et al.* Genetic and pharmacological targeting of activin receptor-like kinase 1 impairs tumor growth and angiogenesis. *J. Exp. Med.* **207**, 85–100 (2010).
- Cunha, S. I. & Pietras, K. ALK1 as an emerging target for antiangiogenic therapy of cancer. *Blood* **117**, 6999–7006 (2011).
- ten Dijke, P. & Arthur, H. M. Extracellular control of TGF β signalling in vascular development and disease. *Nature Rev. Mol. Cell Biol.* **8**, 857–869 (2007).
- Xu, Q. *et al.* Vascular development in the retina and inner ear: control by Norrin and Frizzled-4, a high-affinity ligand-receptor pair. *Cell* **116**, 883–895 (2004).
- Ye, X. *et al.* Norrin, Frizzled-4, and Lrp5 signaling in endothelial cells controls a genetic program for retinal vascularization. *Cell* **139**, 285–298 (2009).

- Hackett, S. F., Wiegand, S., Yancopoulos, G. & Campochiaro, P. A. Angiopoietin-2 plays an important role in retinal angiogenesis. *J. Cell. Physiol.* **192**, 182–187 (2002).
- Haigh, J. J. *et al.* Cortical and retinal defects caused by dosage-dependent reductions in VEGF-A paracrine signaling. *Dev. Biol.* **262**, 225–241 (2003).
- Zhao, J., Sastry, S. M., Sperduto, R. D., Chew, E. Y. & Remaley, N. A. Arteriovenous crossing patterns in branch retinal vein occlusion. The Eye Disease Case-Control Study Group. *Ophthalmology* **100**, 423–428 (1993).
- Kumar, B. *et al.* The distribution of angioarchitectural changes within the vicinity of the arteriovenous crossing in branch retinal vein occlusion. *Ophthalmology* **105**, 424–427 (1998).
- Rakic, J. M. *et al.* Placental growth factor, a member of the VEGF family, contributes to the development of choroidal neovascularization. *Invest. Ophthalmol. Vis. Sci.* **44**, 3186–3193 (2003).
- Takeda, A. *et al.* CCR3 is a target for age-related macular degeneration diagnosis and therapy. *Nature* **460**, 225–230 (2009).
- Van de Veire, S. *et al.* Further pharmacological and genetic evidence for the efficacy of PI3F inhibition in cancer and eye disease. *Cell* **141**, 178–190 (2010).
- Sun, D., Kar, S. & Carr, B. I. Differentially expressed genes in TGF- β 1 sensitive and resistant human hepatoma cells. *Cancer Lett.* **89**, 73–79 (1995).
- Li, X., Miyajima, M., Jiang, C. & Arai, H. Expression of TGF- β s and TGF- β type II receptor in cerebrospinal fluid of patients with idiopathic normal pressure hydrocephalus. *Neurosci. Lett.* **413**, 141–144 (2007).
- Saito, K. *et al.* Gene expression profiling of mucosal addressin cell adhesion molecule-1⁺ high endothelial venule cells (HEV) and identification of a leucine-rich HEV glycoprotein as a HEV marker. *J. Immunol.* **168**, 1050–1059 (2002).
- Spirin, K. S. *et al.* Basement membrane and growth factor gene expression in normal and diabetic human retinas. *Curr. Eye Res.* **18**, 490–499 (1999).
- Gao, B. B., Chen, X., Timothy, N., Aiello, L. P. & Feener, E. P. Characterization of the vitreous proteome in diabetes without diabetic retinopathy and diabetes with proliferative diabetic retinopathy. *J. Proteome Res.* **7**, 2516–2525 (2008).
- Lebrin, F. *et al.* Endoglin promotes endothelial cell proliferation and TGF- β /ALK1 signal transduction. *EMBO J.* **23**, 4018–4028 (2004).
- Anderberg, C. *et al.* Deficiency for endoglin in tumor vasculature weakens the endothelial barrier to metastatic dissemination. *J. Exp. Med.* **210**, 563–579 (2013).
- Mahmoud, M. *et al.* Pathogenesis of arteriovenous malformations in the absence of endoglin. *Circ. Res.* **106**, 1425–1433 (2010).
- Bobik, A. Transforming growth factor- β s and vascular disorders. *Arterioscler. Thromb. Vasc. Biol.* **26**, 1712–1720 (2006).
- ten Dijke, P., Goumans, M. J. & Pardali, E. Endoglin in angiogenesis and vascular diseases. *Angiogenesis* **11**, 79–89 (2008).
- Ray, B. N., Lee, N. Y., How, T. & Blobe, G. C. ALK5 phosphorylation of the endoglin cytoplasmic domain regulates Smad1/5/8 signaling and endothelial cell migration. *Carcinogenesis* **31**, 435–441 (2010).
- Lynch, J. *et al.* MiRNA-335 suppresses neuroblastoma cell invasiveness by direct targeting of multiple genes from the non-canonical TGF- β signalling pathway. *Carcinogenesis* **33**, 976–985 (2012).
- Gregory, A. D., Capoccia, B. J., Woloszynek, J. R. & Link, D. C. Systemic levels of G-CSF and interleukin-6 determine the angiogenic potential of bone marrow resident monocytes. *J. Leukoc. Biol.* **88**, 123–131 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This project was supported by grants from the Lowy Medical Research Foundation, the Medical Research Council, The Wellcome Trust, UCL Business (Proof of Concept Grant) and the Rosetrees Trust. J.W.B.B. is supported by a NIHR Research Professorship. H.M.A. is supported by a British Heart Foundation Senior Fellowship. We would also like to thank M. Gillies for his role in initiating the original project, P. Luthert and C. Thauung for human tissue samples and advice on human pathology specimens, S. Perkins and R. Nan for assistance with the surface plasmon resonance analysis, and P. ten Dijke for discussions and advice.

Author Contributions The project was conceived by J.G., S.E.M. and X.W. Experiments were designed by J.G., S.E.M., X.W. and S.A. Microarrays were performed by J.A.G.M. and qPCR reactions by X.W. X.W. and S.A. characterized the *Lrg1* knockout mice and LRG1 antibody. X.W. performed all the metatarsal assays (except in Fig. 5j, k), aortic ring assays and Matrigel assays, carried out all the biochemical and molecular biology work and analysed the data. S.A. and X.W. undertook the immunohistochemistry and generated the OIR mouse model. U.F.O.L., C.A.K.L., S.A., X.W. and J.W.B.B. performed the CNV experiments, and S.A. and X.W. analysed the data. J.W.B.B. provided human vitreal samples. Z.Z. and H.M.A. generated MLEC;Eng^{fl/fl} cells and X.W. performed proliferation assay and biochemical analysis. Z.Z., S.A. and H.M.A. carried out the metatarsal assays on Eng knockout mice. V.B.T. performed the Biacore experiments. N.J. and M.S. provided assistance and technique support. X.W., S.A., J.G. and S.E.M. produced the figures, and J.G. and S.E.M. wrote the text, with all authors contributing to the final manuscript. J.G. and S.E.M. provided leadership throughout the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.G. (j.greenwood@ucl.ac.uk) or S.E.M. (s.moss@ucl.ac.uk).

METHODS

Animals. C75BL/6J mice were purchased from Harlan Laboratories. *rd1* (ref. 34), *Vldlr*^{-/-} (ref. 35) and *Grhl3*^{cl/j} curly tail mice were purchased from the Jackson Laboratory. *Lrg1*^{-/-} mice were generated by the University of California Davies knockout mouse project (KOMP) repository (<http://www.komp.org/>) and Supplementary Fig. 12). *Rosa26-CreERT;Eng*^{fl/fl} and *Cdh5(PAC)-CreERT2;Eng*^{fl/fl} mice have been previously described^{27,28}. All procedures were performed in accordance with the UK Animals (Scientific Procedures) Act and with the Association for Research in Vision and Ophthalmology Statement for the Use of Animals in Ophthalmic and Vision Research and the Animal Welfare and the Ethical Review Bodies of the UCL Institute of Ophthalmology and Newcastle University.

Vessel isolation and gene expression analysis. Mouse retinal vessels from wild-type C57BL/6J (15 weeks), *Vldlr*^{-/-} (16 weeks), *rd1* (18 weeks) and *Grhl3*^{cl/j} curly tail (13 weeks) mice were isolated as described elsewhere³⁶. RNA was extracted from the enriched microvascular preparations and processed for whole-genome microarray analysis as previously described³⁶. Twelve mice were used per strain per RNA extraction. This was repeated four times, providing RNA for four chips per animal model.

Quantitative PCR (qPCR). RNA was extracted using Trizol (Invitrogen) followed by an RNeasy clean-up (QIAGEN). RNA was reverse transcribed using the QuantiTect Reverse Transcription Kit (QIAGEN) and PCR was conducted with QuantiTect PowerSybr Green (Applied Biosystems) using a 7900HT Fast Real-Time PCR System (Applied Biosystems); samples were normalized to glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*). Primers used in this study are listed in Supplementary Table 2. Student's *t*-test was performed to determine statistical significance between test groups.

SDS-PAGE and western blotting. Proteins were separated by SDS-PAGE. Gels were either stained using Coomassie-blue or transferred onto a Hybond-P PVDF membrane (GE Healthcare). Blots were probed with phospho-Smad1/5 antibody (rabbit monoclonal antibody, NEB), phospho-Smad2 antibody (rabbit monoclonal antibody, NEB), TGF- β 1 antibody (mouse monoclonal, R&D systems), T β R1 antibody (mouse monoclonal, R&D systems), ALK1 antibody (rabbit polyclonal, Santa Cruz Biotechnology), ALK5 antibody (rabbit polyclonal, Abcam), endoglin antibody (mouse monoclonal, R&D Systems), LRG1 antibody (HPA001888, rabbit polyclonal, Sigma) or GAPDH antibody (mouse monoclonal, Novus), followed by horseradish peroxidase (HRP)-conjugated secondary antibodies (GE Healthcare) or HRP-conjugated Protein A (GE Healthcare). Densitometry was performed using ImageJ software (National Institutes of Health). Student's *t*-test was performed to determine statistical significance between test groups.

RNA in situ hybridization. Eyes were fixed in 2% (w/v) paraformaldehyde (PFA) in PBS for 2 min and dissected in 2× PBS. Retinae were flattened and fixed in 100% ice-cold methanol overnight at -20 °C. After recovery from methanol, retinae were re-fixed for 10 min in 4% PFA and washed in PBS before digestion for 10 min in proteinase K (80 μ g ml⁻¹ in PBS). Retinae were re-fixed for 5 min in 4% PFA and 0.2% glutaraldehyde in PBS. After a brief wash in PBS, retinae were pre-incubated in hybridization buffer (50% formamide, 5× SSC, 50 μ g ml⁻¹ transfer RNA, 1% SDS, 50 μ g ml⁻¹ heparin) for 1 h at 65 °C. Denatured RNA probes were incubated with retinae at 65 °C overnight. Primers used to generate RNA probes by PCR are listed in Supplementary Table 2. Probes were labelled with digoxigenin (DIG)-UTP using a DIG RNA labelling kit (Roche Diagnostics). Signal was developed with alkaline phosphatase-conjugated anti-digoxigenin Fab fragments, according to the manufacturer's instructions.

Immunohistochemistry. Retinal/RPE whole-mounts were fixed and stained as previously described³⁷, and incubated overnight with antibodies against human LRG1 (HPA001888), mouse PECAM-1 (rat monoclonal, BD Biosciences), mouse collagen IV (rabbit polyclonal, AbD Serotec), human collagen IV (goat polyclonal, Millipore), rat NG2 (rabbit polyclonal, Millipore), mouse F4/80 (rat monoclonal, AbD Serotec), mouse endoglin (rat monoclonal, Santa Cruz Biotechnology) or human VE-cadherin (mouse monoclonal, Santa Cruz Biotechnology), and identified with Alexa 488, Alexa 594 or Alexa 647 secondary antibodies (Invitrogen) or FITC-GSL (DyLight 594) isolectin B4 (Vector Labs). Retinae were flat-mounted in Mowiol and examined by epifluorescence (Leica DM IRB inverted research microscope or Olympus SZX16 research stereo zoom microscope) or confocal (Carl Zeiss LSM 510 or 710) microscopy. For quantifying the retinal vascular area, raw image data were processed with Photoshop CS4.3. Three-dimensional rendering of confocal Z-stacks was carried out using Imaris 7.5 software (Bitplane AG). The retinal vasculature was analysed through automatic surface rendering aided by manual threshold adjustment so that only blood vessels were included for analysis. Imaris Key Frame Animation was used for movie generation.

Human tissue. Vitreous and plasma samples were collected from patients having surgery for PDR or epiretinal membrane. Human tissue arrays were obtained from Pantomics and stained with rabbit anti-LRG1 antibody (Sigma). The study

followed the ethical guidelines of the Declaration of Helsinki. Institutional Review Boards granted approval for allocation and biochemical analysis of specimens.

Cells and cell culture. Pooled HUVECs were purchased from Lonza and cultured according to suppliers instructions. HEK293T cells were purchased from Invitrogen and cultured as recommended. Mouse primary brain endothelial cells were isolated, purified and cultured as previously described for rat³⁸. The immortalized Lewis rat brain microvascular endothelial cell line GPNT was grown as previously described³⁹. MLECs were isolated from *Rosa26-CreERT;Eng*^{fl/fl} mice carrying the Immortomouse transgene, and were collected and cultured as previously described²⁷. Cells were pre-treated with 1 μ M 4OH-tamoxifen for 48 h in culture to generate ENG-depleted cells (MLEC;*Eng*^{-/-}) and untreated cells served as controls (MLEC;*Eng*^{fl/fl}).

Generation of LRG1 polyclonal antibody. Rabbits were immunized with purified full-length His-tagged human LRG1 protein (Covalab). Pre-immune sera were collected to produce control IgG. Antisera were collected after 3 months and antibody was purified by HiTrap Protein G FF column (GE Healthcare) and concentrated and desalted using HiPrep 26/10 Desalting (GE Healthcare).

Matrigel HUVEC tube formation assay. HUVECs were grown on growth factor-reduced Matrigel (BD Biosciences) as described elsewhere⁴⁰. The 96-well plates were coated with Matrigel-containing diluent (control) or LRG1 (20 μ g ml⁻¹), rabbit polyclonal antibody against LRG1 (C10-54, 100 nM), rabbit IgG (100 nM), ALK1 inhibitor (LDN 193189, Axon Medchem BV, 100 nM) or ALK5 inhibitor (SB43152, Sigma, 10 μ M), and allowed to polymerize in the incubator at 37 °C for 45 min. Tube formation was visualized using an Olympus SZX16 Research stereo-microscope and analysed by counting the number of branch points and total tube length per well using ImageJ. Three independent experiments were carried out and each was performed in triplicate. Student's *t*-test was performed to determine statistical significance between test groups.

Metatarsal angiogenesis assay. The metatarsal angiogenesis assay was carried out as described⁴¹. Metatarsal bones were isolated from E16.5 wild-type control or *Lrg1*^{-/-} littermate mice and treated with TGF- β 1 (5 ng ml⁻¹, R&D systems), LRG1 (20 μ g ml⁻¹), anti-LRG1 polyclonal antibody (100 nM) or rabbit IgG (100 nM) as indicated. Medium was replaced every 2 days. At day 10 of culture, the explants were fixed and stained for PECAM-1 (rat monoclonal, BD Biosciences) and visualized under an Olympus SZX16 Research stereo-zoom microscope. After image processing in Photoshop CS4 to mask the cartilage, the length of PECAM-1-positive tubular structures and the number of branch points were determined by Imaris 7.5 software (Bitplane) using automatic filament tracing with manual threshold corrections. Statistical data were imported into Excel (Microsoft) for calculating total vessel length and the number of branch points. A least three independent experiments were carried out, comprising a minimum of 30 metatarsals for each treatment. Student's *t*-test was performed to determine statistical significance between test groups.

To investigate the effect of ENG depletion on the pro-angiogenic effect of LRG1, metatarsals from *Cdh5(PAC)-CreERT2;Eng*^{fl/fl} mice²⁸ were used to generate endothelial-specific depletion of ENG after addition of 1 μ M 4OH-tamoxifen 3 days after metatarsal bone isolation, when neovessels began to emerge. On day 4, LRG1 was added to a final concentration of 20 μ g ml⁻¹ and the media (including LRG1 and 4OH-tamoxifen supplements) was refreshed every other day until day 12. ENG depletion was confirmed using an anti-ENG antibody (E-Bioscience). Separate experiments using control metatarsals confirmed that 1 μ M 4OH-tamoxifen per se did not affect neovessel formation in the metatarsal angiogenesis assay. Analysis of angiogenesis was carried out as described above. Metatarsals from five independent litters was used and two-way ANOVA was performed to determine statistical significance between test groups.

Aortic ring angiogenesis assay. The aortic ring angiogenesis assay was performed using a modified method described previously⁴². Diameter rings (1 mm) were sliced from aortae of P7 wild-type control or *Lrg1*^{-/-} littermate mice. Aortic rings were then placed in a 96-well plate coated with a rat tail collagen I gel (BD Biosciences) containing LRG1 (20 μ g ml⁻¹), anti-LRG1 polyclonal antibody (100 nM) or rabbit IgG (100 nM) as indicated, and cultured in DMEM supplemented with 2.5% FBS containing relevant compounds. Medium was replaced every 2 days. At day 10 of culture, the explants were fixed, stained for GSL isolectin IB4 (Vector Labs) and visualized under an Olympus SZX16 Research stereo-zoom microscope. The number of sprouts was counted manually. Three independent experiments were carried out with a mean of ≥ 15 aortic rings being analysed for each treatment. Student's *t*-test was performed to determine statistical significance between test groups.

Mouse model of CNV. CNV was induced as described elsewhere^{36,43}. In the anti-LRG1 antibody blockade, study animals received an intravitreal injection of immediately after the laser burn. Antibody at a concentration of 1, 2.5, 5 or 10 μ g (each in 1 μ l) of the anti-LRG1 polyclonal antibody was delivered to one eye and a pre-immune IgG, serving as control, delivered to the contralateral eye.

Five (in the case of antibody treatment) or seven days after injury, CNV lesions were imaged as described before^{36,43,44}. Mice were then killed for retina and RPE flat-mount preparation and mRNA extraction. The CNV lesions were visualized after FITC-conjugated GSL isolectin B4 (Vector Labs) and mouse PECAM-1 (rat polyclonal, BD Biosciences) staining using an Olympus SZX16 Research stereozoom microscope and a Zeiss LSM 710 confocal microscope. Student's *t*-test was performed to determine the statistical significance between wild-type and *Lrg1* knockout mice. One-way ANOVA was used to test statistical significance between antibody treatment groups.

Mouse model of OIR. Nursing mothers and neonatal mice were placed in a 75% oxygen supply chamber from P7 to P12 and exposed to a standard 12 h light–dark cycle as previously described⁴⁵. The extent of vaso-obliteration was determined in retinal flat-mounts at P12, and the extent of normal vessel regrowth and neovascularization were evaluated at P17 as previously described⁴⁶. Retinae were also recovered for mRNA extraction and analysis at P12 and P17. The effect of antibody blockade on retinal revascularisation and neovascular tuft formation was carried out by delivering anti-LRG1 blocking antibody (50 mg kg^{−1} intraperitoneal in 100 µl at P13 and P15), anti-VEGFR2 blocking antibody (DC101, 12.5 mg kg^{−1} intraperitoneal at P13 and P15) or a combination of the two, followed by assessment of the vasculature at P17. Student's *t*-test was performed to determine the statistical significance between wild-type and *Lrg1* knockout mice. One-way ANOVA was used to test statistical significance between antibody treatment groups.

Co-immunoprecipitation. Primary mouse brain endothelial cells from wild-type or *Lrg1*^{−/−} littermate mice, GPNT cells or HUVECs were lysed in RIPA buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate and 1% Nonidet P-40). Soluble peptide-tagged extracellular domains of TβRII (Myc-tagged), ALK1 (HA-tagged), ALK5 (HA-tagged) and ENG (V5-tagged), as well as full-length LRG1 (His-tagged), were generated in separate cultures of transfected HEK293T cells, and serum-free media containing the individual proteins was collected after 5 days. Non-tagged secreted extracellular domains of the TGF-β receptors were also generated in an identical manner. Media containing individual extracellular domains of a TGF-β receptor were incubated with media containing LRG1 in the presence or absence of TGF-β1 at 4 °C with rotation before immunoprecipitation. After pre-clearing, cell lysates or recombinant protein mixtures were incubated with TGF-β receptor antibodies or anti-LRG1 antibody-conjugated protein G beads at 4 °C overnight and then fractionated by SDS-PAGE and blotted. The membranes were probed with antisera as described earlier.

Proliferation assay. Mouse primary brain endothelial cells from *Lrg1*^{−/−} and wild-type mice were cultured in EGM2 media supplemented with puromycin (5 µg ml^{−1}) until sub-confluent, followed by 48 h serum starvation in EBM2 medium. Cells were stimulated with TGF-β1 (5 ng ml^{−1}), LRG1 (20 µg ml^{−1}) or TGF-β1 plus LRG1 in EBM2 medium at 37 °C. After 3 h, cells were fixed and stained with an antibody to Ki67 (mouse monoclonal antibody, Dako) to detect proliferating cells. The proliferation rate was evaluated as the percentage of Ki67-positive cells of the total endothelial cell number per well. Student's *t*-test was performed to determine the statistical significance between treatment groups.

Molecular biological methods. The coding sequence of human *LRG1* (NM_052972) carrying a 6×His tag or HA tag at the 3' end and Kozak consensus sequence at the 5' end was cloned into pcDNA3.1 (Invitrogen) at the HindIII/XhoI sites to form pcDNA-LRG1-His or pcDNA-LRG1-HA (Supplementary Fig. 8). The coding sequence of human *LRG1* was cloned into a pEGX4T1 GST expression vector (GE Healthcare) at the BamHI/SalI site to form glutathione S-transferase (GST)–ENG. The extracellular domain of human TβRII (NM_001024847.2) carrying a Myc tag, ALK1 (NM_000020.2) carrying a HA tag, ALK5 (NM_004612.2) carrying a HA tag, ENG (NM_001114753.1) carrying a V5 tag at the 3' end and Kozak consensus sequence at the 5' end were cloned into pcDNA3.1 at HindIII/EcoRI sites. The recombinant human proteins were expressed in HEK293T cells (Invitrogen). siRNA oligonucleotides (SASI_Rn01_00111211 (Sigma)) were used for *Lrg1* gene knockdown in GPNT cells, and siRNA oligonucleotides (ON-TARGETplus SMARTpools, Thermo Scientific) were used for knockdown in HUVECs of *ALK1* (L-005302-00-0005), *ALK5* (L-003929-00-0005), *ENG* (L-011026-00-0005) and *TβRII* (L-001000-00-50), and control siRNA (D-001810-10-05) was used as a negative control for knockdown in HUVECs.

Lipofectamine 2000 transfection reagent (Invitrogen) was used for transfection of mammalian cells. Oligofectamine 2000 transfection reagent (Invitrogen) was used for siRNA knockdown in GPNT cells. GeneFECTOR transfection reagent

(VennNova) was used for siRNA knockdown in HUVECs. PCR and qPCR primer sequences are shown in Supplementary Table 2.

Purification of recombinant proteins. LRG1–His was expressed in HEK293T cells and purified using HisPrep FF16/10 column (GE Healthcare) and buffer exchanged into PBS using HiPrep 26/10 Desalting (GE Healthcare) according to manufacturer's instruction. GST–ENG was expressed in BL21-competent cells and purified using glutathione Sepharose 4B (GE Healthcare) and eluted in elution buffer (50 mM Tris-HCl, 10 mM reduced glutathione, pH 8.0) according to the manufacturer's instruction. Denatured LRG1–His protein was generated by boiling at 100 °C for 15 min.

Surface plasmon resonance. All surface plasmon resonance experiments were carried out on a BiacoreT200 instrument (GE Healthcare). LRG1 was covalently immobilized via primary amino groups on a CMS sensor chip as per manufacturer's instructions (specific contact time 20 s at a flow rate of 10 µl min^{−1}; LRG1 concentration at 25 µg ml^{−1} diluted using 10 mM sodium acetate, pH 5.0). The amount of immobilized LRG1 corresponded to 2,000 response units in flow cell 2. Flow cell 1 on the same sensor chip, reserved for control runs, was treated identically but without LRG1 immobilization. For all SPR measurements, GST-tagged ENG was diluted in running buffer (1× PBS, pH 7.2). The association was monitored by injecting different concentrations (1–50 nM) of the analyte (ENG) into channels 1 and 2, starting with the lowest analyte concentration. All experiments were conducted in triplicates at 25 °C at a flow rate of 30 µl min^{−1}. The association time for ligand–analyte steady state binding was optimised to 180 s and a subsequent 300 s were allowed for dissociation. Between injections the sensor chip surface was regenerated with glycine-HCl, pH 2.0, at a flow rate of 30 µl min^{−1} for 30 s. All curves were corrected for nonspecific binding by subtraction of control curves obtained from injection of the analyte through the blank flow cell 1. The affinity and dissociation constants were calculated from the plots of the steady-state binding as a function of protein concentration, using the Biacore T200 evaluation software and a homogenous 1:1 Langmuir binding kinetic model. The analysis provided values for the dissociation affinity constant (*K_D*), the association rate constant (*K_a*) and the dissociation rate constant (*K_d*). **Statistical analyses.** Data are represented as mean ± s.e.m. Statistical analyses were performed by Student's *t*-test or one-way ANOVA followed by Tukey/Bonferroni post-test analysis or two-way ANOVA as appropriate, using Prism 5 (GraphPAD Software Inc.). **P* < 0.05; ***P* < 0.01; ****P* < 0.001. Each represents significant statistical comparisons among the listed (*x* axis) experimental groups.

34. Blanks, J. C. & Johnson, L. V. Vascular atrophy in the retinal degenerative *rd* mouse. *J. Comp. Neurol.* **254**, 543–553 (1986).
35. Heckenlively, J. R. *et al.* Mouse model of subretinal neovascularization with choroidal anastomosis. *Retina* **23**, 518–522 (2003).
36. McKenzie, J. A. *et al.* Apelin is required for non-neovascular remodelling in the retina. *Am. J. Pathol.* **108**, 399–409 (2012).
37. Fruttiger, M. Development of the mouse retinal vasculature: angiogenesis versus vasculogenesis. *Invest. Ophthalmol. Vis. Sci.* **43**, 522–527 (2002).
38. Abbott, N. J., Hughes, C. C., Revest, P. A. & Greenwood, J. Development and characterisation of a rat brain capillary endothelial culture: towards an in vitro blood–brain barrier. *J. Cell Sci.* **103**, 23–37 (1992).
39. Romero, I. A. *et al.* Changes in cytoskeletal and tight junctional proteins correlate with decreased permeability induced by dexamethasone in cultured rat brain endothelial cells. *Neurosci. Lett.* **344**, 112–116 (2003).
40. Arnaoutova, I. & Kleinman, H. K. *In vitro* angiogenesis: endothelial cell tube formation on gelled basement membrane extract. *Nature Protocols* **5**, 628–635 (2010).
41. Deckers, M. *et al.* Effect of angiogenic and antiangiogenic compounds on the outgrowth of capillary structures from fetal mouse bone explants. *Lab. Invest.* **81**, 5–15 (2001).
42. Nicosia, R. F. & Ottinetti, A. Growth of microvessels in serum-free matrix culture of rat aorta. A quantitative assay of angiogenesis in vitro. *Lab. Invest.* **63**, 115–122 (1990).
43. Balagun, K. S. *et al.* ElAV vector-mediated delivery of endostatin or angiostatin inhibits angiogenesis and vascular hyperpermeability in experimental CNV. *Gene Ther.* **13**, 1153–1165 (2006).
44. Toma, H. S., Barnett, J. M., Penn, J. S. & Kim, S. J. Improved assessment of laser-induced choroidal neovascularization. *Microvasc. Res.* **80**, 295–302 (2010).
45. Smith, L. E. *et al.* Oxygen-induced retinopathy in the mouse. *Invest. Ophthalmol. Vis. Sci.* **35**, 101–111 (1994).
46. Connor, K. M. *et al.* Quantification of oxygen-induced retinopathy in the mouse: a model of vessel loss, vessel regrowth and pathological angiogenesis. *Nature Protocols* **4**, 1565–1573 (2009).

Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR

Daniel H. Ebert^{1,2}, Harrison W. Gabel¹, Nathaniel D. Robinson¹, Nathaniel R. Kastan¹, Linda S. Hu¹, Sonia Cohen¹, Adrijia J. Navarro¹, Matthew J. Lyst³, Robert Ekiert³, Adrian P. Bird³ & Michael E. Greenberg¹

Rett syndrome (RTT) is an X-linked human neurodevelopmental disorder with features of autism and severe neurological dysfunction in females. RTT is caused by mutations in methyl-CpG-binding protein 2 (MeCP2), a nuclear protein that, in neurons, regulates transcription, is expressed at high levels similar to that of histones, and binds to methylated cytosines broadly across the genome^{1–5}. By phosphotryptic mapping, we identify three sites (S86, S274 and T308) of activity-dependent MeCP2 phosphorylation. Phosphorylation of these sites is differentially induced by neuronal activity, brain-derived neurotrophic factor, or agents that elevate the intracellular level of 3',5'-cyclic AMP (cAMP), indicating that MeCP2 may function as an epigenetic regulator of gene expression that integrates diverse signals from the environment. Here we show that the phosphorylation of T308 blocks the interaction of the repressor domain of MeCP2 with the nuclear receptor co-repressor (NCoR) complex and suppresses the ability of MeCP2 to repress transcription. In knock-in mice bearing the common human RTT missense mutation R306C, neuronal activity fails to induce MeCP2 T308 phosphorylation, suggesting that the loss of T308 phosphorylation might contribute to RTT. Consistent with this possibility, the mutation of MeCP2 T308A in mice leads to a decrease in the induction of a subset of activity-regulated genes and to RTT-like symptoms. These findings indicate that the activity-dependent phosphorylation of MeCP2 at T308 regulates the interaction of MeCP2 with the NCoR complex, and that RTT in humans may be due, in part, to the loss of activity-dependent MeCP2 T308 phosphorylation and a disruption of the phosphorylation-regulated interaction of MeCP2 with the NCoR complex.

The location of RTT missense mutations provides insight into MeCP2's function. Of the four most common RTT missense mutations, three—at R106, R133 and T158—are in the DNA-binding domain and disrupt binding to methylated DNA, suggesting that binding to methylated DNA is critical for MeCP2 function^{6–8}. Disruption of binding to methylated DNA probably impairs MeCP2's function as a repressor. Consistent with this possibility, the fourth common RTT missense mutation, R306C, is located within the repressor domain of MeCP2. However, the mechanism of action of the MeCP2 repressor domain and the specific functions of R306 were not known.

Recent evidence indicates that sensory stimulation triggers MeCP2 phosphorylation at a specific site, S421, raising the possibility that MeCP2 may function as a neuronal activity-regulated repressor, and that RTT may result from the de-regulation of neuronal activity-dependent gene programs^{9–13}. However, studies of knock-in mice in which S421 is converted to an alanine have challenged this hypothesis, as this mutation had no detectable effect on gene transcription¹⁴.

To search for additional activity-dependent sites of MeCP2 phosphorylation that might regulate MeCP2 function, we performed phosphotryptic mapping of MeCP2 derived from ³²P-orthophosphate-labelled neurons that were left untreated or exposed to elevated levels of KCl to trigger membrane depolarization and calcium influx. Lysates from these

neurons were incubated with anti-MeCP2 antibodies, and immunoprecipitates were resolved by SDS-PAGE. The band corresponding to MeCP2 was excised and digested with trypsin. Phosphopeptides were resolved by two-dimensional thin-layer electrophoresis and chromatography. Autoradiography of the phosphotryptic maps revealed a complex pattern of MeCP2 phosphorylation in both untreated and membrane-depolarized neurons, indicating that MeCP2 is phosphorylated at many sites in cultured neurons (Fig. 1a). However, three phosphopeptides, indicated as *a*, *b* and *c* in Fig. 1a, appeared reproducibly after membrane depolarization. The same inducible phosphopeptides were detected in MeCP2 S421A knock-in neurons, indicating that these phosphopeptides do not contain S421.

To identify the site(s) of inducible MeCP2 phosphorylation, we compared phosphotryptic maps of MeCP2 phosphorylated *in vitro* by calcium-regulated kinases with the phosphotryptic maps of MeCP2 obtained from membrane-depolarized neurons. Once a kinase was identified that phosphorylated MeCP2 *in vitro* at a site that co-migrated with spots *a*, *b*, or *c* on the phosphotryptic map from primary neuronal culture, we mutated MeCP2 to identify the candidate sites of phosphorylation. To characterize further these sites of MeCP2 phosphorylation, we generated phosphorylation site-specific antibodies to each of the sites. This analysis (Fig. 1 and Supplementary Figs 1–6) revealed that upon membrane depolarization, or upon stimulation with the GABA_A (γ-aminobutyric acid subtype A)-receptor antagonist bicuculline, which relieves inhibitory input and allows for the release of endogenous glutamate in the cultures, MeCP2 becomes newly phosphorylated at S86, S274, T308 and S421. We note that S86 and T308 phosphorylation was not detected by previous mass spectrometry studies, underscoring the value of using phosphotryptic mapping to discover sites of activity-dependent phosphorylation in neurons.

To investigate whether phosphorylation of these sites on MeCP2 is inducible *in vivo*, mice were treated with kainic acid to trigger seizures and robust neuronal activity. Forebrain lysates from untreated and kainic-acid-injected mice were analysed by western blotting. We found that exposure to kainic acid reproducibly induced MeCP2 phosphorylation at S86, S274, T308 and S421 (Fig. 1b). In brain lysates from mice not exposed to kainic acid, a low level of immune reactivity is detected, suggesting that basal activity in the brain also induces phosphorylation of MeCP2 at each of these sites. These findings demonstrate that phosphorylation at MeCP2 S86, S274, T308 and S421 is induced by neuronal activity, both in cell culture and in the intact brain.

We next compared the ability of different extracellular stimuli to induce the phosphorylation of MeCP2. Cortical neurons were stimulated with KCl to induce membrane depolarization, with brain-derived neurotrophic factor (BDNF), or with forskolin to activate protein kinase A (PKA) (Fig. 1d). Western blotting of lysates of these stimulated cultures revealed that MeCP2 phosphorylation at S86 and S274 is induced significantly by either BDNF or forskolin and less well upon membrane depolarization with KCl. By contrast, MeCP2 phosphorylation at T308 and S421 is induced most effectively by membrane

¹Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA. ²Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA. ³Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, UK.

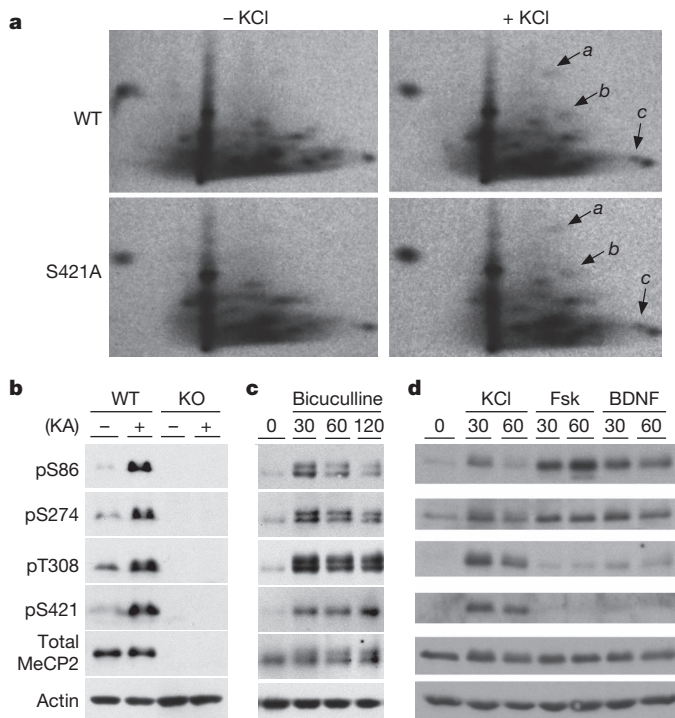


Figure 1 | Phosphotryptic mapping of MeCP2 identifies activity-dependent phosphorylation sites. **a**, Dissociated cortical cultures, derived from wild-type (WT) or MeCP2 S421A knock-in mice, were metabolically labelled with ^{32}P -orthophosphate and were left untreated or membrane depolarized with KCl. Phospho-peptides were spotted in the lower left corner of the plate, resolved by thin-layer electrophoresis in the horizontal direction and thin-layer chromatography in the vertical direction. The activity-induced spots are indicated (arrows a–c). **b**, Mice were injected with kainic acid (KA) to induce seizures. **c**, Cortical neurons were stimulated with bicuculline. **d**, Cortical neurons were membrane depolarized with KCl or treated with BDNF or forskolin (Fsk). Time of stimulation (min) is indicated. Lysates were examined by western blotting with the indicated antibodies. Phospho-site-specific antibodies are indicated by 'p' followed by the site. Findings were confirmed with at least three biological replicates from dissections of independent animals.

depolarization and less potently by BDNF or forskolin. These findings indicate that MeCP2 may be a convergence point in the nucleus for multiple signalling pathways and raise the possibility that differential phosphorylation of MeCP2, bound broadly across the genome, could mediate the response of neuronal chromatin to diverse stimuli. In a manner similar to the epigenetic regulation of gene expression by modifications of histones, the multiple stimulus-regulated post-translational modifications of MeCP2 may be a mechanism that modulates chromatin remodelling in post-mitotic neurons.

To assess the importance of phosphorylation at these novel sites for neuronal function and RTT, we focused our attention on the phosphorylation of MeCP2 T308 because of its proximity to common RTT missense mutations R306C/H. A possible clue to the function of phosphorylation of MeCP2 T308 was provided by a recent study demonstrating that the R306C mutation disrupts the ability of MeCP2 to interact with the NCoR complex⁸. NCoR forms a complex with multiple proteins, including histone deacetylase 3 (HDAC3), and this complex is thought to trigger histone deacetylation and gene repression^{15–17}. Given the proximity of T308 to amino acids that are essential for recruitment of the NCoR complex, we postulated that phosphorylation of MeCP2 at T308 might affect the interaction of MeCP2 with the NCoR complex and might thereby mediate activity-dependent changes in gene expression.

We developed a peptide pull-down assay to examine the interaction of the repressor domain of MeCP2 with the NCoR complex and assessed the effect of MeCP2 T308 phosphorylation on this interaction (Fig. 2a, b and

Supplementary Figs 7–9). We synthesized biotin-conjugated MeCP2-derived peptides in which T308 was either left unphosphorylated (np peptide) or phosphorylated at T308 (pT308 peptide), mixed the peptides with streptavidin-conjugated magnetic beads, and, by western blotting with various antibodies to components of the NCoR complex, assessed the ability of the beads to pull down the NCoR complex from brain lysates. The np peptide was able to pull down core components of the NCoR complex including HDAC3, TBL1, TBLR1 and GPS2, but not another co-repressor SIN3A, indicating that the region of MeCP2 surrounding T308 contains a binding site that specifically mediates the interaction of MeCP2 with the NCoR complex. By contrast, the pT308 peptide did not interact at all with the NCoR complex. Similarly, peptides containing phosphomimetic T308D and T308E mutations, acidic amino acid mutations that partially mimic a phosphorylated residue, showed reduced binding to the NCoR complex. The peptide pull-down experiments demonstrate that the carboxy-terminal region of MeCP2's transcription repression domain interacts with the NCoR complex and that phosphorylation of T308 abrogates this interaction. These findings suggest that neuronal-activity-induced phosphorylation of MeCP2 T308 disrupts the interaction of the repressor domain of MeCP2 with the NCoR complex and raise the possibility that, by altering the interaction of NCoR with MeCP2, the phosphorylation of T308 might affect MeCP2-dependent transcription. However, it remains to be determined whether the phosphorylation of MeCP2 T308 leads to a complete release of the NCoR complex from MeCP2 bound to methylated DNA, or if T308 phosphorylation disrupts the interaction of the MeCP2 repressor domain with NCoR without leading to a release of the NCoR complex.

To determine if MeCP2 T308 phosphorylation affects MeCP2's function as a transcriptional repressor, we assessed the ability of wild-type and mutant versions of MeCP2 (R306C, T308A, T308D and T308E) to repress reporter gene transcription (Fig. 2c and Supplementary Figs 8 and 9). Cultured cortical neurons were co-transfected with plasmid constructs expressing MeCP2 variants fused to the GAL4 DNA-binding domain (GAL4–MeCP2) and a firefly luciferase reporter plasmid with a promoter containing GAL4-binding sites (luciferase reporter)⁸. Upon transfection into cortical neurons together with the luciferase reporter, wild-type GAL4–MeCP2 effectively represses reporter gene transcription. However, insertion of the R306 to cysteine

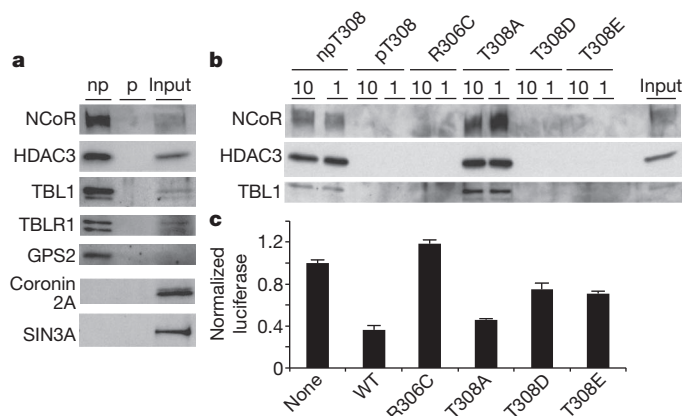


Figure 2 | Phosphorylation of T308 regulates MeCP2's interaction with the NCoR complex. **a**, Peptides corresponding to MeCP2, either not phosphorylated (np) or synthesized to be phosphorylated at the amino acid residue corresponding to 308 (p), were incubated with cortical neuron lysates. Pull downs were visualized by western blotting with the indicated antibodies. **b**, Peptide pull downs using synthetic peptides (at 10× or 1× concentration) containing one of the following mutations: R306C, T308A, T308D, or T308E. **c**, Luciferase reporter gene assay with plasmids encoding reporter and GAL4–MeCP2 variants, as indicated, transfected into cortical neurons. 'None' had no GAL4–MeCP2 variant transfected. Shown are averages and s.e.m. of three biological replicates. Findings for each experiment were confirmed with at least three biological replicates from dissections of independent litters.

mutation into GAL4–MeCP2 resulted in a version of GAL4–MeCP2 that is no longer capable of repressing transcription. Given that the mutation of MeCP2 R306 to C renders MeCP2 incapable of binding the NCoR complex, these findings suggest that GAL4–MeCP2 represses the GAL4 reporter gene in an NCoR-dependent manner. If the phosphorylation of MeCP2 at T308 blocks the ability of MeCP2 to repress transcription via the NCoR complex, we would expect that mutation of T308 to an acidic amino acid (D or E), which partially abolishes the interaction of MeCP2 with the NCoR complex, should partially suppress GAL4–MeCP2-dependent transcription repression. This is what we observed when GAL4–MeCP2 T308D or GAL4–MeCP2 T308E were tested for their ability to repress reporter gene transcription. The intermediate loss of the transcription repression by the GAL4–MeCP2 T308D/E variants corresponds with partial loss of binding to the NCoR complex that we observed by western blotting (Supplementary Fig. 9). By contrast, GAL–MeCP2 T308A, a mutant MeCP2 that is still capable of interacting with the NCoR complex, was fully capable of repressing luciferase reporter gene transcription. These findings suggest that phosphorylation of MeCP2 T308 prevents the interaction of the repressor domain of MeCP2 with the NCoR complex, thereby reducing MeCP2–NCoR–HDAC3-mediated transcriptional repression.

We next asked if the activity-dependent phosphorylation of MeCP2 T308 affects the ability of MeCP2 to function as a repressor of activity-dependent gene transcription. Towards this end we generated mice in which MeCP2 T308 is converted to an alanine (MeCP2 T308A knock-in mice) and assessed the effect of this mutation on activity-dependent gene transcription. We first demonstrated by western blotting that MeCP2 T308A knock-in mice and their wild-type littermates express equivalent levels of MeCP2 protein (Supplementary Fig. 10a, b). This indicates that the T308A mutation does not alter the stability of MeCP2. In addition, we confirmed by western blotting with anti-MeCP2 phospho-T308 antibodies that the MeCP2 T308A knock-in neurons lack T308 phosphorylation (Supplementary Fig. 10c). We also demonstrated by chromatin immunoprecipitation with anti-MeCP2 antibodies that the T308A mutation does not affect MeCP2 binding to DNA (Supplementary Fig. 10d), and by peptide pull-down experiments (Fig. 2b) and co-immunoprecipitation of MeCP2 and NCoR from forebrain extracts (Supplementary Fig. 10e) that the T308A mutation does not disrupt the basal binding of MeCP2 to the NCoR complex. These findings suggest that any abnormality that we detect in gene transcription in MeCP2 T308A knock-in mice might be attributed to the loss of the phosphorylation-dependence of the interaction of MeCP2 with the NCoR complex rather than to a decrease in MeCP2's expression, binding to DNA, or ability to interact with NCoR.

We assessed the effect of the MeCP2 T308A mutation on activity-dependent gene transcription directly by exposing cultured neurons derived from wild-type and MeCP2 T308A knock-in mice to elevated levels of KCl and monitoring activity-dependent gene expression by PCR with reverse transcription (RT–PCR) (Fig. 3a). We found that membrane depolarization induces *Arc*, *Fos*, *Nptx2* and *Adcyap1* messenger RNA expression equivalently in wild-type and MeCP2 T308A knock-in neurons, indicating that the signalling apparatus that conveys the membrane depolarization/calcium signal to the nucleus to activate gene transcription functions normally in MeCP2 T308A knock-in neurons. By contrast, membrane depolarization induces significantly less *Npas4* in MeCP2 T308A knock-in neurons than in wild-type neurons. Previous studies have shown that *Npas4* expression is induced upon membrane depolarization of excitatory neurons and that NPAS4 promotes the development of inhibitory synapses on excitatory neurons¹⁸, a process that has been found to be abnormal in RTT¹⁹. NPAS4 is a transcription factor that has been suggested to regulate inhibitory synapse number by activating expression of *Bdnf* (ref. 18). Therefore, we investigated whether *Bdnf* might also be impaired in T308A knock-in neurons compared to wild-type neurons. There is a trend towards decreased induction of *Bdnf* mRNA in T308A knock-in neurons compared to wild-type neurons. We also observed an attenuation of light

induction of *Npas4* and *Bdnf* in the visual cortex of dark-reared T308A knock-in compared to wild-type mice but no statistically significant difference in *Arc*, *Fos*, *Nptx2* and *Adcyap1* mRNA transcription in these two strains of mice (Fig. 3b). This suggests that the decrease in activity-dependent *Npas4* and *Bdnf* expression in T308A knock-in compared to wild-type mice occurs *in vivo* and could, in principle, contribute to neural circuit defects that occur in RTT. These findings are consistent with a model in which activity-dependent phosphorylation of MeCP2 T308 leads to a decrease in the association of the NCoR co-repressor complex with the repressor domain of MeCP2, thus facilitating activity-dependent *Npas4* transcription and the subsequent activation of *Bdnf* transcription. However, given that MeCP2 binds broadly across the genome, we cannot rule out the possibility that, in MeCP2 T308A knock-in mice, the reduction in neuronal activity-dependent induction of *Npas4* and *Bdnf* mRNA is due to an effect of the T308A mutation on chromatin architecture that affects excitatory/inhibitory balance and only indirectly leads to a reduction in the levels of *Npas4* and *Bdnf* mRNA.

Finally, we sought to determine if the disruption of activity-dependent phosphorylation of MeCP2 T308 and the consequent disruption of activity-dependent gene transcription contributes to RTT. We first noted that T308 is in close proximity to common RTT missense mutations at R306C/H. Given that the kinases that can phosphorylate T308 (CaMKIV and PKA) typically require a basophilic residue two or three amino acids amino-terminal to the site of phosphorylation²⁰, we hypothesized that R306C/H mutations, in addition to abolishing the interaction of MeCP2 with the NCoR complex, might render MeCP2 refractory to phosphorylation at T308. To test this hypothesis, we

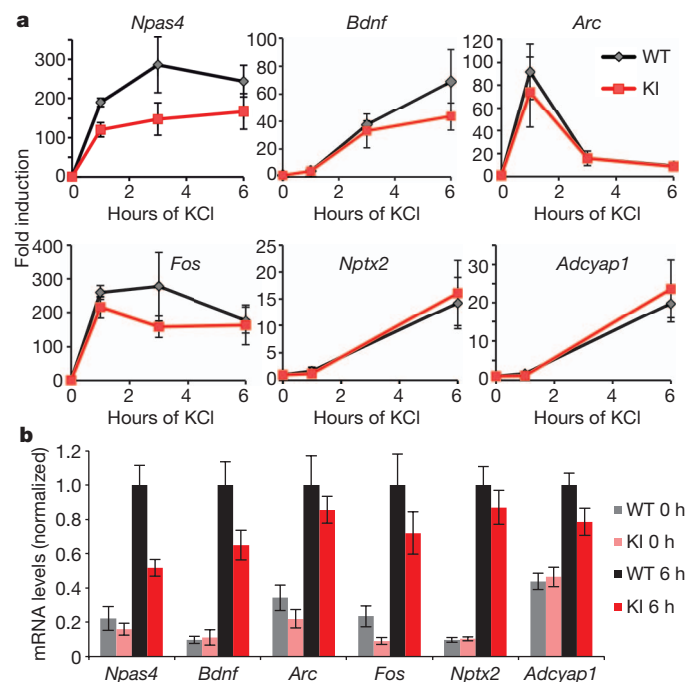


Figure 3 | MeCP2 T308A knock-in mice have altered activity-dependent gene expression. **a**, Dissociated cortical neuron cultures, generated from MeCP2 T308A knock-in mice (KI, red) and wild-type littermates (WT, black), were membrane depolarized with KCl. Shown are fold-inductions of RT–PCR values for the average of three independent days of dissection (biological replicates). For *Npas4*, *P* values were 0.024 (repeated measures ANOVA) and 0.03 at 1 h and 0.01 at 6 h (paired two-tailed *t*-test). **b**, Eight-week-old mice, kept in the dark for the previous 2 weeks, were either kept in the dark (0 h) or exposed to a light source for 6 h (6 h). From dissected visual cortices, RT–PCR values were divided by the average of WT 6 h. The numbers of littermate mice averaged per time point were: WT 0 h (*n* = 6), KI 0 h (*n* = 6), WT 6 h (*n* = 8) and KI 6 h (*n* = 10). Error bars indicate s.e.m. For *Npas4* at 6 h, the two-tailed *t*-test *P* value is 0.0007 and for *Bdnf* at 6 h, the *P* value is 0.04.

exposed wild-type or MeCP2 R306C knock-in mice⁸ to kainic acid, prepared lysates from hippocampi, and assessed the phosphorylation of MeCP2 at T308 by western blotting (Fig. 4a). Exposure of mice to kainic acid induced the phosphorylation of MeCP2 T308 in wild-type but not MeCP2 R306C knock-in mice despite equivalent expression of total MeCP2 in both genotypes. Notably, we confirmed that the anti-MeCP2 pT308 antibodies are still able to recognize phosphorylated-T308 in the presence of the R306C mutation (Supplementary Fig. 11). Taken together, these findings indicate that the common R306C/H mutations that occur in RTT not only disrupt the interaction of MeCP2 with the NCoR, they also abrogate activity-dependent phosphorylation of MeCP2 at T308. Thus, RTT in individuals with R306C/H mutations could result simply from the loss of basal NCoR binding to MeCP2, which, by necessity, would abolish the regulated interaction of MeCP2 with NCoR. However, it is possible that the loss of activity-dependent MeCP2 T308 phosphorylation could, in and of itself, contribute to aspects of RTT in these individuals. It is also possible that the loss of MeCP2 T308 phosphorylation could have consequences, in addition to the disruption of the proper regulation of NCoR binding, which may also be relevant to the aetiology of RTT.

To investigate whether activity-dependent MeCP2 T308 phosphorylation might contribute to RTT, we asked if MeCP2 T308A knock-in mice display neurological impairments that are hallmarks of RTT, including reduced brain weight, motor abnormalities, and a reduced threshold for the onset of seizures (Fig. 4b–e and Supplementary Fig. 12). As discussed above, MeCP2 T308A knock-in mice, when compared to wild-type littermates, have normal levels of MeCP2 protein expression, binding to DNA, and interaction with the NCoR complex. These findings suggest that any neurological phenotypes observed in the MeCP2 T308A knock-in mice are most probably due to the disruption of T308 phosphorylation and the loss of the phosphorylation-dependence of the interaction of MeCP2 with the NCoR complex. The first indication that MeCP2 T308A knock-in mice have neurological deficits was that the brains of MeCP2 T308A knock-in mice weigh

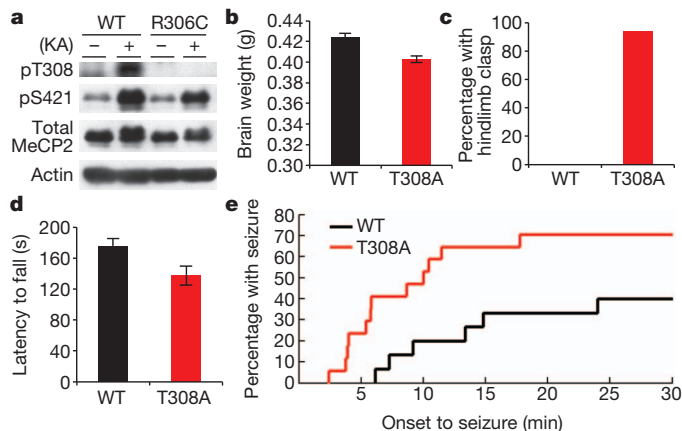


Figure 4 | Phosphorylation of MeCP2 T308 is a key mechanism underlying RTT. **a**, MeCP2 R306C knock-in mice and wild-type littermates (WT) were treated with kainic acid (KA) to induce seizures or left untreated (–). Forebrain lysates were resolved by western blot analysis with indicated antibodies. Findings were confirmed with five biological replicates from independent animals. **b**, Brains were dissected and weighed for T308A knock-in mice ($n = 16$) and wild-type littermates ($n = 13$) with 5% reduction in the knock-in and P value between genotypes of 0.00006 (unpaired two-tailed t -test). **c**, T308A knock-in mice ($n = 16$) and wild-type littermates ($n = 13$) were scored for the presence of a hindlimb clasp. P value was 0.00005 (two-proportion Z -test). **d**, T308A knock-in mice ($n = 16$) and wild-type littermates ($n = 13$) were placed on an accelerating rotarod, and latency to fall was measured with a P value between genotypes of 0.02 (two-tailed t -test). **e**, T308A knock-in mice ($n = 17$) and wild-type littermates ($n = 15$) were injected with 40 mg kg^{–1} PTZ and latency to onset of generalized tonic-clonic seizures was measured. P value was 0.029 (two-sample Kolmogorov–Smirnov test). Error bars indicate s.e.m.

significantly less than the brains of their wild-type littermates despite the fact that the overall body weights of these two types of mice are similar (Fig. 4b). We also found that when compared to wild-type littermate controls, MeCP2 T308A knock-in mice display hindlimb claspings and a reduced ability to stay on an accelerating rotarod, two phenotypes that indicate that MeCP2 T308A knock-in mice have motor system defects (Fig. 4c, d). To determine whether MeCP2 T308A knock-in mice have a lower seizure threshold, wild-type and MeCP2 T308A knock-in mice were exposed to a low-dose of the GABA antagonist pentylenetetrazol (PTZ), and the time to onset and frequency of generalized tonic-clonic seizures measured (Fig. 4e). Compared to wild-type littermates, the MeCP2 T308A knock-in mice have more seizures, and the onset of the seizures occurs more rapidly. These findings suggest that the MeCP2 T308A knock-in mice have a lower seizure threshold compared to wild-type mice. This decrease in seizure threshold could be due to the decrease in *Npas4* and *Bdnf* transcription in MeCP2 T308A knock-in mice and the consequent disruption of excitatory/inhibitory balance in the brains of these animals^{18,21}. Although a direct comparison has not yet been performed, the MeCP2 R306C knock-in mice clearly have a more severe phenotype than the MeCP2 T308A knock-in mice⁸, consistent with the R306C mutation abolishing the binding to the NCoR complex and the T308A mutation disrupting the activity-regulated interaction with the NCoR complex. Taken together, these findings indicate that the loss of activity-regulated phosphorylation of T308, and the disruption of activity-dependent control of the interaction of MeCP2 with the NCoR complex, probably contribute to some of the neurological deficits in RTT.

How could loss of NCoR binding (MeCP2 R306C mice⁸) and constitutive NCoR binding (MeCP2 T308A mice) both lead to a RTT-like syndrome? A possible answer may come from previous studies demonstrating that both loss of MeCP2 and overexpression of MeCP2 can lead to RTT-like symptoms, although of varying severity^{22,23}. The R306C phenotype may be analogous to MeCP2 loss of function RTT (MeCP2 can no longer bind NCoR), whereas the T308A phenotype may be similar to MeCP2 gain-of-function phenotype (MeCP2 constitutively binds NCoR and is a constitutively active repressor). Taken together, the MeCP2 R306C and MeCP2 T308A knock-in studies provide evidence that the interaction of MeCP2 with the NCoR complex is critical for proper MeCP2 function, and that dysregulation of this interaction can lead to RTT.

METHODS SUMMARY

Phosphorylation mapping was performed on MeCP2 derived from neurons that were pre-treated with ³²P-orthophosphate and membrane-depolarized with KCl. MeCP2 T308A knock-in mice were generated as described previously¹⁴.

Full Methods and any associated references are available in the online version of the paper.

Received 25 July 2012; accepted 4 June 2013.

Published online 16 June 2013.

1. Nan, X., Campoy, F. J. & Bird, A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell* **88**, 471–481 (1997).
2. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).
3. Skene, P. J. *et al.* Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol. Cell* **37**, 457–468 (2010).
4. Amir, R. E. *et al.* Rett syndrome is caused by mutations in X-linked MeCP2, encoding methyl-CpG-binding protein 2. *Nature Genet.* **23**, 185–188 (1999).
5. Chahrour, M. *et al.* MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**, 1224–1229 (2008).
6. Ballestar, E., Yusufzai, T. M. & Wolffe, A. P. Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry* **39**, 7100–7106 (2000).
7. Ho, K. L. *et al.* MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell* **29**, 525–531 (2008).
8. Lyst, M. J. *et al.* Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nature Neurosci.* <http://dx.doi.org/10.1038/nn.3434> (16 June 2013).
9. Chen, W. G. *et al.* Derepression of *Bdnf* transcription involves calcium-dependent phosphorylation of MeCP2. *Science* **302**, 885–889 (2003).

10. Zhou, Z. *et al.* Brain-specific phosphorylation of Mecp2 regulates activity-dependent Bdnf transcription, dendritic growth, and spine maturation. *Neuron* **52**, 255–269 (2006).
11. Tao, J. *et al.* Phosphorylation of Mecp2 at serine 80 regulates its chromatin association and neurological function. *Proc. Natl Acad. Sci. USA* **106**, 4882–4887 (2009).
12. Li, H., Zhong, X., Chau, K. F., Williams, E. C. & Chang, Q. Loss of activity-induced phosphorylation of Mecp2 enhances synaptogenesis, Ltp and spatial memory. *Nature Neurosci.* **14**, 1001–1008 (2011).
13. Ebert, D. H. & Greenberg, M. E. Activity-dependent neuronal signalling and autism spectrum disorder. *Nature* **493**, 327–337 (2013).
14. Cohen, S. *et al.* Genome-wide activity-dependent Mecp2 phosphorylation regulates nervous system development and function. *Neuron* **72**, 72–85 (2011).
15. Kokura, K. *et al.* The Ski protein family is required for Mecp2-mediated transcriptional repression. *J. Biol. Chem.* **276**, 34115–34121 (2001).
16. Li, J. *et al.* Both corepressor proteins Smrt and N-CoR exist in large protein complexes containing Hdac3. *EMBO J.* **19**, 4342–4350 (2000).
17. Perissi, V., Jepsen, K., Glass, C. K. & Rosenfeld, M. G. Deconstructing repression: evolving models of co-repressor action. *Nature Rev. Genet.* **11**, 109–123 (2010).
18. Lin, Y. *et al.* Activity-dependent regulation of inhibitory synapse development by Npas4. *Nature* **455**, 1198–1204 (2008).
19. Dani, V. S. *et al.* Reduced cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett syndrome. *Proc. Natl Acad. Sci. USA* **102**, 12560–12565 (2005).
20. Wayman, G. A., Lee, Y. S., Tokumitsu, H., Silva, A. J. & Soderling, T. R. Calmodulin-kinases: modulators of neuronal development and plasticity. *Neuron* **59**, 914–931 (2008).
21. Hong, E. J., McCord, A. E. & Greenberg, M. E. A biological function for the neuronal activity-dependent component of Bdnf transcription in the development of cortical inhibition. *Neuron* **60**, 610–624 (2008).
22. Collins, A. L. *et al.* Mild overexpression of Mecp2 causes a progressive neurological disorder in mice. *Hum. Mol. Genet.* **13**, 2679–2689 (2004).
23. Ramocki, M. B., Tavyev, Y. J. & Peters, S. U. The Mecp2 duplication syndrome. *Am. J. Med. Genet. A.* **152A**, 1079–1088 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by NIH grant R01NS048276 and the Rett Syndrome Research Trust to M.E.G. D.H.E. was supported by NIH grant K08MH90306, the Dupont–Warren Fellowship in the Department of Psychiatry at Harvard Medical School, and the Nancy Lurie Marks Fellowship in Autism at Harvard Medical School. H.W.G. was supported by Damon Runyon Cancer Research Foundation Grant DRG-2048-10. The Mouse Gene Manipulation Facility of the Boston Children's Hospital Intellectual and Developmental Disabilities Research Center (IDDR), funded by NIH grant P30-HD 18655, assisted in generation of the knock-in mice. We thank members of the Greenberg laboratory, particularly C. Mandel-Brehm and E. Griffith, and also G. Mandel and R. S. Greenberg for helpful discussions.

Author Contributions D.H.E. and M.E.G. conceived and designed the experiments and wrote the manuscript. D.H.E. performed or directed all the experiments in the manuscript. D.H.E., N.D.R. and N.R.K. generated and characterized the MeCP2 T308A knock-in and R306C knock-in mice and characterized activity-dependent phosphorylation of MeCP2. H.W.G. performed the ChIP analysis, and H.W.G. and S.C. performed experiments investigating activity-dependent phosphorylation of MeCP2 that informed this study. D.H.E., L.S.H. and N.D.R. developed the phospho-site-specific antibodies, and A.J.N. assisted in early work with these antibodies during summer rotations. M.J.L., R.E. and A.P.B. discovered that the RTT missense mutation R306C disrupted both the interaction with NCoR and MeCP2's ability to provide transcription repression using the luciferase reporter assay. All authors reviewed the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.E.G. (michael_greenberg@hms.harvard.edu).

METHODS

Gene nomenclature. To maintain consistency of nomenclature with past descriptions of phosphorylation of MeCP2 S421 and RTT missense mutations, the S86, S274, T308 and S421 nomenclature refers to the mouse MeCP2 isoform 2 (MeCP2_e2; NCBI reference sequence NP_034918). S86, S274, T308 and S421 in mouse MeCP2 isoform 2 correspond to S103, S291, T325 and S438, respectively, in the mouse MeCP2 isoform 1 (MeCP2_e1; NCBI reference sequence NP_001075448), correspond to S86, S274, T308 and S423 in the human MeCP2 isoform 1 (NCBI reference sequence NP_004983), and correspond to S98, S286, T320 and S435 in human MeCP2 isoform 2 (NCBI reference sequence NP_001104262). Alternative splicing generates the two MeCP2 isoforms, which are distinguished by distinct N-terminal sequences.

Neuronal cell culture. Primary dissociated cortical neuron cultures were derived from cortices of mice at embryonic day 16 (E16), as previously described²⁴, and cultured for varying days *in vitro* (DIV), allowing for neuronal maturation and synapse development in culture. Cortical cultures were maintained in neurobasal medium with B27 supplement (Invitrogen), 1 mM L-glutamine, and 100 U ml⁻¹ penicillin/streptomycin. Cells were plated at 1–2 × 10⁶ on 6-well dishes, 10 × 10⁶ on 10-cm dishes, and 30 × 10⁶ on 15-cm dishes that had been pre-treated with polyornithine.

Phosphotryptic mapping of MeCP2. Dissociated E16 mouse cortical neurons at 6 DIV were treated overnight with 1 μM tetrodotoxin and 100 μM APV (Tocris Bioscience) to reduce endogenous neuronal activity in the culture. At 7 DIV, cortical neuron cultures, in 10-cm dishes, were labelled with 2.5 mCi of ³²P-orthophosphate (Perkin Elmer) in phosphate-free neurobasal medium for 5 h. Neurons were then left untreated or exposed to 55 mM KCl by addition of 0.5 volumes of depolarization buffer (170 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂ and 10 mM HEPES, pH 7.5) for 90 min, to induce membrane depolarization and robustly model neuronal activity. Neurons were washed once in PBS, lysed in TTN lysis buffer (30 mM Tris, pH 7.5, 1 M NaCl, 3% Triton X-100, 5 mM EDTA, 10 mM β-glycerolphosphate, 10 mM NaF, 2 mM Na₃VO₄, and 1× complete EDTA-free protease inhibitor cocktail (Roche)), and sheared with a 22-gauge needle. Insoluble material was pelleted at 17,000g and removed. Lysates were diluted with equal volumes of H₂O to reduce NaCl concentration to 500 mM. Lysates were immunoprecipitated with an anti-total MeCP2 antibody (antibody to the C terminus of MeCP2 that was generated in-house as described previously¹⁰) bound to protein A sepharose beads for 2 h while rotating at 4 °C. Immunoprecipitates were washed four times in TTN buffer (diluted to 500 mM NaCl and 1.5% Triton X-100) and resolved by SDS–PAGE. The dried gel was exposed to autoradiography. Phosphotryptic mapping of MeCP2 followed the procedure detailed in ref. 25. The MeCP2 band was excised from the gel and digested with trypsin (TPCK-treated, Worthington). The tryptic phosphopeptides were separated in two-dimensions by thin-layer electrophoresis, using pH 1.9 electrophoresis buffer (2.5% formic acid (88% w/v) and 7.8% glacial acetic acid), for 30 min at 1,000 V and by thin-layer chromatography, using the phosphochromatography buffer (37.5% *n*-butanol, 25% pyridine, and 7.5% glacial acetic acid), on glass-backed TLC plates (20 × 20 cm, 100 μM cellulose, EM Science). The phosphotryptic maps were visualized by autoradiography. Experiments shown were repeated greater than three times using biological replicates from dissection of independent litters.

In vitro kinase assays. MeCP2 fragments were generated by calcium phosphate transfection of HEK 293T cells with constructs expressing Flag-tagged N-terminal MeCP2 variants from amino acid 1 to 193 or C-terminal MeCP2 variants from amino acid 173 to 484. Missense mutations at putative sites of phosphorylation were generated by site-directed mutagenesis using Quickchange (Stratagene) and fully sequenced through the entire subcloned segment. HEK 293T cells were transfected by calcium phosphate. Forty-eight hours after transfection, the exogenous MeCP2 variants were collected in lysis buffer (50 mM Tris, pH 7.5, 500 mM NaCl, 2.5% Triton X-100, 2 mM EDTA, 10 mM NaF, 2 mM Na₃VO₄, 1 mM DTT, and 1× complete EDTA-free protease inhibitor cocktail (Roche)), immunoprecipitated with anti-Flag antibodies (M2, Sigma), and eluted from the beads with Flag peptides at 150 ng μl⁻¹ concentration. The purified MeCP2 variants were phosphorylated using *in vitro* kinase assays. For *in vitro* kinase assays with CaMKIV, C-terminal fragments of MeCP2 were incubated in a reaction mixture with 40 mM Tris, pH 7.5, 10 mM MgCl₂, 0.5 mM CaCl₂, 1 mM DTT, 50 μg ml⁻¹ calmodulin (Calbiochem), purified CaMKIV (recombinant, *Escherichia coli*, Life Technologies), 0.1 mM cold ATP, and 5 μCi (0.033 μM) [γ-³²P]ATP (Perkin Elmer) in a 25 μl reaction for 10–30 min at 30 °C. For *in vitro* kinase assays with PKA, purified MeCP2 variants were incubated in a reaction mixture with 40 mM Tris, pH 7.5, 10 mM MgCl₂, 1 mM DTT, PKA (catalytic subunit, mouse, recombinant, *E. coli*, Calbiochem), 0.1 mM cold ATP and 5 μCi (0.033 μM) [γ-³²P]ATP in a 25 μl reaction for 10–30 min at 30 °C.

Generation of anti-MeCP2 phospho-site-specific antibodies. The polyclonal antibody that specifically recognizes S86-phosphorylated MeCP2 was generated by injecting New Zealand White rabbits (Covance Research Products) with the

peptide KQRR(pS)IIRDRGPM-C (Tufts Synthesis Facility) conjugated to KLH. The antiserum was affinity-purified by incubation with a column that was conjugated with phosphorylated-S86 MeCP2 peptide, and the affinity-purified antibody was eluted. This eluate was then incubated with a column conjugated with unphosphorylated-S86 MeCP2 peptide, and the affinity-purified anti-MeCP2 pS86 antibody was collected in the flow-through. The polyclonal antibody that specifically recognizes S274-phosphorylated MeCP2 was generated by injecting rabbits with the peptide RKPG(pS)VVA AAAAEAKKK-C conjugated to KLH. The antibody was affinity purified similar to the purification of the anti-MeCP2 pS86 antibodies. The polyclonal antibody that specifically recognizes T308-phosphorylated MeCP2 was generated by injecting rabbits with the peptide C-TVLPIKKRK(pT)RE conjugated to KLH. The antibody was purified over a column conjugated with MeCP2 T308 peptide, and the affinity-purified anti-MeCP2 pT308 was eluted. The generation of the polyclonal rabbit antibody that specifically recognizes S421-phosphorylated MeCP2 and the polyclonal antibody that recognizes total MeCP2 irrespective of phosphorylation status were previously described¹⁰.

Stimulation of MeCP2 phosphorylation in cell culture and *in vivo*. Cortical neuron cultures (E16 + 7 DIV) were membrane depolarized with 55 mM KCl by addition of 0.5 volumes of depolarization buffer (170 mM KCl, 2 mM CaCl₂, 1 mM MgCl₂ and 10 mM HEPES, pH 7.5). Alternatively, cultures were treated with 20 μM forskolin (Calbiochem) or 50 ng ml⁻¹ BDNF (Peprotech) for 30 min or 1 h. For bicuculline experiments, E16 + 14 DIV cortical neuron cultures were treated with 20 μM bicuculline (Sigma) for 30–120 min. For western blot analysis, cells were lysed in boiling sample buffer to preserve endogenous phosphorylation events and prevent spurious phosphorylation events after cell lysis. Lysates were boiled for 10 min, passed through Wizard Minicolumns (Promega) to remove larger molecules and insoluble material, and resolved by 8% SDS–PAGE gels, normalized by cell number. Western blotting was performed with antibodies specific to MeCP2 phosphorylation sites (generated in our laboratory as described above) or specific to total MeCP2 irrespective of phosphorylation status (Men-8, Sigma) or β-actin (ab8226, Abcam), all at 1:1,000 dilutions. Western blotting was completed with HRP-conjugated secondary antibodies and enhanced chemiluminescence. Seizures were induced in adult C57B/6 male mice, or in MeCP2 knockout male mice (MeCP2^{tm1.1Bird} line acquired from Jackson laboratories), 8–10 weeks of age, by intraperitoneal injection of kainic acid at a dosage of 25 mg kg⁻¹. Ninety minutes after injection, forebrains were collected and lysed in boiling sample buffer to preserve phosphorylation sites, and lysates were analysed by western blotting as described previously. Experiments shown were replicated at least three times using biological replicates from dissection of independent animals with the same results.

Endogenous co-immunoprecipitation. Eight-week-old C57B/J mice were left untreated or seizures were induced by intraperitoneal injection of kainic acid at a dosage of 25 mg kg⁻¹. Two hours after injection, forebrains were isolated and lysed in NP-40 lysis buffer (10 mM HEPES, pH 7.9, 3 mM MgCl₂, 10 mM KCl, 10 mM NaF, 1 mM Na₃VO₄, 0.5 mM DTT, 0.5% NP-40, 1× complete EDTA-free protease inhibitor cocktail (Roche)), dounced 15× with a tight pestle, and pelleted at 1,000g. Lysates were diluted 1:1 with benzonase buffer (10 mM HEPES, pH 7.9, 3 mM MgCl₂, 280 mM NaCl, 0.2 mM EDTA, 10 mM NaF, 1 mM Na₃VO₄, 0.5 mM DTT, 0.5% NP-40, and 1× complete EDTA-free protease inhibitor cocktail (Roche)) and digested with 250 units of the permissive nuclease benzonase (Novagen) for 1 h rotating at 4 °C to release MeCP2 and its protein binding partners from the genome. Digested lysates were pelleted at 17,000g for 20 min at 4 °C and immunoprecipitated with anti-total MeCP2 antibodies (raised in house as in ref. 10), in either the presence of 150 mM NaCl or 250 mM NaCl as indicated, for 2 h while rotating at 4 °C. The peptide-block control was immunoprecipitation of lysates with anti-total MeCP2 antibodies in the presence of the peptide to which the antibody was raised. Western blots of SDS–PAGE resolved immunoprecipitates are shown using anti-NCoR (PA1-844A, Pierce) and anti-MeCP2 antibodies (generated in-house).

Peptide pull-down assays. To investigate the impact of phosphorylation of MeCP2 T308 on binding to other proteins, we synthesized a peptide corresponding to MeCP2 amino acids 285–319 with biotin conjugated to the N terminus of the peptide (Tufts University Core Facility). This peptide (Biotin-KKAVKESSIR SVHETVLPKKRK[T]RETVSIEVKEV) was left unphosphorylated or phosphorylated at the amino acid residue corresponding to T308 (the bold T in brackets). We synthesized additional variant peptides, including a variant in which the amino acid residue corresponding to R306 was synthesized as a cysteine (R306C) and variants in which the amino acid residue corresponding to T308 was synthesized as either an alanine, glutamic acid, or aspartic acid (T308A, T308D, or T308E). The biotin-conjugated peptides, from 0.1 to 5 μg, were bound to 40 μl Streptavidin MagneSphere Paramagnetic Particles (Promega) during a greater than 1 h rotation at 4 °C before incubation with neuronal cell lysates. Cortical neuron cultures (E16 + 7 DIV) were scraped in PBS, lysed in lysis buffer (10 mM HEPES, pH 7.8, 500 mM NaCl, 1% Triton X-100, 10 mM NaF, 1 mM Na₃VO₄, 5 mM EDTA, 0.5 mM DTT,

and 1× complete EDTA-free protease inhibitor cocktail (Roche)), sheared with a 22-gauge needle, and pelleted at 17,000g for 20 min at 4 °C. Lysates were diluted 1:2 with dilution buffer (10 mM HEPES, pH 7.8, 0.5% Triton X-100, 10 mM NaF, 1 mM Na₃VO₄, 5 mM EDTA, 0.5 mM DTT, and 1× complete EDTA-free protease inhibitor cocktail (Roche)) to lead to a final NaCl concentration in the lysate of approximately 167 mM. Neuronal lysates were incubated with biotin-conjugated beads bound to streptavidin particles, rotating at 4 °C for 16 h. The peptide pull-down was washed four times with wash buffer (10 mM HEPES pH 7.8, 150 mM NaCl, 0.5% Triton X-100, 10 mM NaF, 1 mM Na₃VO₄, 5 mM EDTA, and 0.5 mM DTT) and boiled in 1.2× sample buffer for 10 min. The peptide pull-downs were resolved by SDS-PAGE, 5% gel for NCoR and 10% gels for the other proteins. Western blotting was performed with antibodies specific to NCoR (PA1-844A, Pierce), HDAC3 (H3034, Sigma), TBL1 (H-367, Santa Cruz), TBLR1 (ab13799, Abcam), GPS2 (H-225, Santa Cruz), Coronin 2A (M-105, Santa Cruz) and SIN3A (N-19, Santa Cruz).

Transcription repression domain assay. Cortical neurons (2.5 × 10⁵) were plated into wells of 24-well plate and neurons (E16 + 5 DIV) were transfected by calcium phosphate procedure²⁶ with multiple plasmids. All cells were transfected with a plasmid encoding the firefly luciferase with constitutively active TK-promoter and 5× UAS binding sites for GAL4 upstream of the promoter and with a plasmid encoding renilla luciferase, to normalize transfection between samples^{1,8}. In the different conditions, cells were transfected with plasmids encoding fusion proteins between GAL4 and MeCP2 variants. The GAL4–MeCP2 fusion proteins extend from MeCP2 amino acid 201 to 484 and are missing the DNA-binding domain from MeCP2. The GAL4–MeCP2 fusion proteins are brought to the reporter plasmid by interaction between the UAS sequence and GAL4 subunit of the fusion protein. We generated missense mutations at amino acid residue corresponding to MeCP2 T308 to A, D, and E by site-directed mutagenesis using Quikchange (Stratagene). Neurons were transfected with 0.4 µg firefly luciferase plasmid, 0.08 µg renilla luciferase plasmid, 0.03 µg GAL4–MeCP2 variant plasmid, and PCS2 filler plasmid to 1 µg total per well. At 7 DIV, cells were collected for measurement of firefly and renilla luciferase, using Dual-Luciferase Reporter Assay (Promega). The ratio of firefly to renilla luciferase was calculated for each well. Shown is average of ratio of firefly to renilla luciferase of three biological replicates with standard deviation, normalized to condition without any GAL4–MeCP2 variant transfected (labelled ‘none’ in Fig. 2d). The experiment was repeated independently three times from dissections of independent litters with the same results.

MeCP2 R306C knock-in mice. MeCP2 R306C knock-in mice were generated in the Greenberg laboratory as previously described¹⁴, and the homologous recombination was confirmed by sequencing and Southern blot analysis. Eight-week-old MeCP2 R306C knock-in male mice and wild-type littermates were injected with 25 mg kg^{−1} of kainic acid to induce seizures. After one hour, forebrains were dissected and lysed in boiling sample buffer and with shearing from a Polytron. Lysates were resolved for western blot analysis with the anti-MeCP2 pT308, anti-MeCP2 pS421, and anti-total MeCP2 antibodies. The experiment was repeated with five biological replicates from independent animals with the same results.

MeCP2 T308A knock-in mice. MeCP2 T308A knock-in mice were generated using the same approach as previously described¹⁴, and the homologous recombination was confirmed by sequencing and Southern blot analysis. The targeting construct contained the mutation, ACC to GCC, for the codon corresponding to amino acid 308. Given that MeCP2 is on the X chromosome, all experiments used male knock-in or wild-type littermates.

To determine whether MeCP2 T308A knock-in mice could be phosphorylated at T308, 10-week-old mice were injected with 25 mg kg^{−1} of kainic acid, or left untreated, and, after one hour, forebrain lysates were resolved for western blot analysis with indicated antibodies. For MeCP2 protein levels, brains were dissected from MeCP2 T308A knock-in mice and wild-type littermates. Brains were lysed in boiling sample buffer, sheared with a Polytron, and resolved for western blot analysis.

For MeCP2 ChIP, forebrains of 11-week-old mice were dissected on ice, and crosslinking and nuclear preps were performed as described¹⁴. To fragment chromatin, SDS was added to 0.3% final concentration, and samples were sonicated using a Covaris S2 sonicator (12 min, 5% duty cycle, power level 4, 200 cycles per burst). Similar fragmentation was confirmed for all samples by gel electrophoresis, with ~80% of the DNA appearing as a smear from approximately 100–500 bp in length. For immunoprecipitation, rabbit polyclonal antisera recognizing total MeCP2⁹ were used as previously described¹⁴ except that ChIP buffer containing SDS instead of NaDOC was used for the sample during the immunoprecipitation (10 mM Tris pH 8.0, 0.1% SDS, 1% Triton X-100, 150 mM NaCl, 1 mM EDTA, 0.3 mM EGTA, 1× Roche complete EDTA-free protease inhibitors, 10 mM β-glycerolphosphate, 10 mM NaF). After overnight incubation the supernatant was discarded and beads were washed at 4 °C with the following washes: 2 times with low salt wash buffer (0.1% SDS, 20 mM Tris pH 8.0, 1% Triton X-100, 150 mM

NaCl, 2 mM EDTA), 2 times with high salt wash buffer (0.1% SDS, 20 mM Tris pH 8.0, 1% Triton X-100, 500 mM NaCl, 2 mM EDTA), 2 times with LiCl wash buffer (0.1% NaDOC, 10 mM Tris pH 8.0, 1% NP40, 250 mM LiCl, 1 mM EDTA), once with TE. DNA was eluted by incubation for 30 min at 65 °C in TE containing 1% SDS. Input and immunoprecipitated DNA were decrosslinked by incubating for 12–16 h at 65 °C, treated with RNase (20 µM RNase A at 37 °C for 0.5–1 h) and proteinase K (280 µM at 55 °C for 2 h), phenol/chloroform extracted twice, chloroform extracted once, and DNA was isolated using a Qiagen PCR purification column (Qiagen).

Quantitative PCR analysis was carried out using the StepOnePlus qPCR system and Power SYBR Green mix (Life technologies). The ‘fraction of input’ value for each amplicon (primers listed below) was determined by comparing the average threshold cycle of the immunoprecipitated DNA to a standard curve generated using serial dilutions of the input DNA and interpolating the ‘fraction of input’ value for the sample. All sites of amplification showed significant >10 fold increased signal compared to MeCP2 ChIP done from MeCP2 knockout mice. Primers used for ChIP-qPCR, identical to those used previously¹⁴, were *Bdnf* upstream, GGCCAAGTGAATTGGGTAT, TGATGGCAGCAATGTTTCTC, −29 kb to TSS; *Actb* TSS, AGTGTCTACACCGCGGAAT, CTGGCAGACCCA ACTTTACG, 236 bp to TSS; *Npas4* TSS, AGGGACCCAGGTTTTCAT, GGGCTTCAGACCACCCTAAT, −313 bp to TSS; and major satellite, GGCG AGAAACTGAAAATCACG, AGGTCCTTCAGTGTGCATTTC.

For gene expression analysis in the stimulated visual cortex, MeCP2 T308A knock-in male mice and wild-type male littermates were placed in the dark at 6 weeks of age. Two weeks later, mice were kept in the dark or exposed to light for 6 h. Visual cortices were dissected. RNA was purified by TRIzol (Life Technologies) extraction, RNeasy Mini kit (Qiagen) with on-column DNase digestion (RNase-Free DNase Set, Qiagen). From 1.0 µg RNA, cDNA was generated using random primers (High Capacity cDNA Reverse Transcription Kit, Applied Biosystems). Quantitative PCR (qPCR) was performed using LightCycler 480 Real-Time PCR instrument (Roche) with LightCycler 480 SYBR Green 1 master mix (Roche). The primers used for *Npas4* were GCTATACTCAGAAGGTCCA GAAGGC, TCAGAGATGAGGGTAGCACAGC; *Bdnf*, GATGCCGCAAAA CATGCTCTATGA, TAATACTGTACACACGCTCAGCTC; *Arc*, TACCGTTA GCCCTATGCCCATC, TGATATTGCTGAGCCTCAACTG; β-tubulin, CGAC AATGAAGCCCTCTACGAC, ATGGTGGCAGACACAAGGTGGTTG.

Values at each time point were normalized to β-tubulin. To illustrate the induced gene expression on one graph, values were divided by average of the wild-type 6-h time point for each gene tested. Sample size was chosen to detect magnitude of gene expression changes consistent with magnitude of gene expression changes reported in MeCP2 knockout mice. The *P* values were calculated by unpaired two-tailed Student's *t*-test.

In addition, dissociated E16.5 cortical neuron cultures were generated from MeCP2 T308A knock-in males and wild-type littermates and 7.5 × 10⁵ cells were plated per well of a 6-well dish. Cultures were fed at 7 DIV with 30% fresh NB media. At 10 DIV, cultures were treated with AP5 and TTX to silence activity in the culture for 2 h before starting the 6-h membrane depolarization time point. Cultures were membrane depolarized with 55 mM KCl for 1 h, 3 h, or 6 h or left untreated. Cells were lysed in TRIzol, and RNA purified and cDNA generated as above. Three wells per condition in an experiment were combined to make one sample. To show fold-induction of gene expression over the time course, values at each time point were divided by the value at 0 h. Three independent days of dissection and experiments (biological replicates) were averaged for the qPCR experiments shown. *P* values were calculated by two-way repeated measures ANOVA and by two-tailed Student's *t*-test at specific time points, pairing wild-type and knock-in neurons derived from the same litters, combined on the day of dissection, for the three independent days of dissection and culturing of neurons.

MeCP2 T308A knock-in males (*n* = 16) and wild-type littermate males (*n* = 13) were weighed at 14–16 weeks of age. Whole brains were then dissected and weighed. A second independent cohort of MeCP2 T308A knock-in mice (*n* = 9) and wild-type littermates (*n* = 9) had identical findings with the same magnitude of difference between genotypes. *P* values were calculated by two-tailed, unpaired student's *t*-test.

To determine the presence of a hindlimb clasp, MeCP2 T308A knock-in male mice (*n* = 16) and wild-type male mice (*n* = 13), at 11–13 weeks of age, were lifted by their tails, to a height one foot off the table. The presence of a hindlimb clasp was defined as pulling in one or both of the hindlimbs fully towards the body for at least 2 s. Each mouse was scored, blinded to genotype, for the presence of a hindlimb clasp during three rounds of two-minute observation, with 5 min between each round. A second independent cohort of MeCP2 T308A knock-in mice (*n* = 9) and wild-type littermates (*n* = 9) had identical findings. *P* value, calculated by two-proportion *Z*-test, was 0.00005.

MeCP2 T308A knock-in mice ($n = 16$) and wild-type littermates ($n = 13$) were tested on an accelerating rotarod (Economex, Columbus Instruments) at 13–15 weeks of age. Animals were brought in 30 min before testing for habituation. Each animal was placed on an accelerating rotarod set to 4.0–40 r.p.m. over a period of 5 min. A fall was called either when an animal fell off the rod or rotated twice around without recovery. Littermates were given two to four trials with an hour of rest in between, and an average latency to fall was calculated for each animal. The statistical test used to compare fall latency across the two genotypes was a two-tailed, unpaired Student's *t*-test.

To evaluate seizure threshold, MeCP2 T308A knock-in mice ($n = 17$) and wild-type littermates ($n = 15$) were injected with pentylenetetrazol (PTZ), a GABA receptor antagonist, at 14–16 weeks of age. Mice were habituated to the room for 20 min and weighed. Mice were injected intraperitoneally with 40 mg kg^{-1} of PTZ (Sigma Aldrich). Mice were scored for time to onset of a generalized tonic-clonic seizure for 30 min following injection of PTZ.

The behavioural characterization of the T308A knock-in mice in this manuscript was performed at fifth generation backcross to C57B/6 from the 129J ES cell line used to generate the mice. Experiments involving mice were performed blinded to genotype. Sample size for behavioural experiments, of 13–17 male mice per genotype, was chosen to mitigate against genetic background variance. Only litters with at least one male of each genotype, T308A knock-in and wild type, were

used for analysis. All mice in the behavioural experiments had the same tests and experiences; there was no randomization used. Mice were tested in the following order: hindlimb clasp, rotarod and PTZ-induction of seizures. There was at least one week between tests. The independent two-tailed *t*-tests used met the test criteria in that the samples were independent, data in each sample were independent, and all population values appear normally distributed (unimodal histogram and symmetric). For the PTZ-induced seizures, a two-sample Kolmogorov–Smirnov (KS2) test was used to determine whether two one-dimensional probability distributions differ. Variances across genotypes for all tests appear homoscedastic, as variances of s.d. are similar. All animal experiments were in compliance with ethical regulations and were approved by the Harvard Medical Area Standing Committee on Animals (HMA IACUC).

24. Xia, Z., Dudek, H., Miranti, C. K. & Greenberg, M. E. Calcium influx via the NMDA receptor induces immediate early gene transcription by a MAP kinase/ERK-dependent mechanism. *J. Neurosci.* **16**, 5425–5436 (1996).
25. Meisenhelder, J., Hunter, T. & van der Geer, P. Phosphopeptide mapping and identification of phosphorylation sites. *Curr. Protoc. Mol. Biol.* Ch. 18, Unit 18 19 (2001).
26. Dudek, H., Ghosh, A. & Greenberg, M. E. Calcium phosphate transfection of DNA into neurons in primary culture. *Curr. Protoc. Neurosci.* Ch. 3, Unit 3 11 (2001).

Promoter directionality is controlled by U1 snRNP and polyadenylation signals

Albert E. Almada^{1,2*}, Xuebing Wu^{1,3*}, Andrea J. Kriz², Christopher B. Burge^{2,3} & Phillip A. Sharp^{1,2}

Transcription of the mammalian genome is pervasive, but productive transcription outside of protein-coding genes is limited by unknown mechanisms¹. In particular, although RNA polymerase II (RNAPII) initiates divergently from most active gene promoters, productive elongation occurs primarily in the sense-coding direction^{2–4}. Here we show in mouse embryonic stem cells that asymmetric sequence determinants flanking gene transcription start sites control promoter directionality by regulating promoter-proximal cleavage and polyadenylation. We find that upstream antisense RNAs are cleaved and polyadenylated at poly(A) sites (PASs) shortly after initiation. *De novo* motif analysis shows PAS signals and U1 small nuclear ribonucleoprotein (snRNP) recognition sites to be the most depleted and enriched sequences, respectively, in the sense direction relative to the upstream antisense direction. These U1 snRNP sites and PAS sites are progressively gained and lost, respectively, at the 5' end of coding genes during vertebrate evolution. Functional disruption of U1 snRNP activity results in a dramatic increase in promoter-proximal cleavage events in the sense direction with slight increases in the antisense direction. These data suggest that a U1–PAS axis characterized by low U1 snRNP recognition and a high density of PASs in the upstream antisense region reinforces promoter directionality by promoting early termination in upstream antisense regions, whereas proximal sense PAS signals are suppressed by U1 snRNP. We propose that the U1–PAS axis limits pervasive transcription throughout the genome.

Two potential mechanisms for suppressing transcription elongation in the upstream antisense region of gene transcription start sites (TSSs) include inefficient release of paused RNAPII and/or early termination of transcription. RNAPII pauses shortly after initiation downstream of the gene TSS and the paused state is released by the recruitment and activity of the positive transcription elongation factor b (P-TEFb)⁵. A detailed characterization of several upstream antisense RNAs (uaRNAs) in mouse embryonic stem cells (ESCs) suggested that P-TEFb is recruited similarly in both sense and antisense directions⁶, and in human cells, elongating RNAPII (phosphorylated at Ser 2 in the carboxy-terminal domain) occupies the proximal upstream transcribed region⁷. These data suggest that the upstream antisense RNAPII complex undergoes the initial phase of elongation but probably terminates early owing to an unknown mechanism.

To test globally whether upstream antisense transcripts undergo early termination (compared to coding messenger RNA) by a canonical PAS-dependent cleavage mechanism, we mapped by deep sequencing the 3' ends of polyadenylated RNAs in mouse ESCs (N. Spies, C.B.B. & D. P. Bartel, unpublished protocol). For most protein-coding genes, transcription termination is triggered by cleavage of the nascent RNA upon recognition of a PAS, whose most essential feature is an AAUAAA sequence or a close variant located about 10–30 nucleotides upstream of the cleavage site⁸. We sequenced two complementary DNA (cDNA) libraries and obtained over 230 million reads, of which 114 million mapped uniquely to the genome with at most two mismatches.

We developed a computational pipeline to identify 835,942 unique 3' ends (cleavage sites) whose poly(A) tails are likely to be added post-transcriptionally and are also associated with the canonical PAS hexamer or its common variants (Supplementary Fig. 1, see Methods).

To investigate whether uaRNAs are terminated by PAS-dependent mechanisms, we focused our analysis on cleavage sites proximal to gene TSSs and at least 5 kilobases (kb) away from known gene transcription end sites (TESs). Interestingly, in the upstream antisense region we observed a twofold higher number of cleavage sites compared to the downstream sense sites flanking protein-coding-gene TSSs (Fig. 1a). The peak of the upstream antisense cleavage sites is about 700 bases from the coding-gene TSS. This observation suggests that upstream antisense transcripts are frequently terminated by PAS-directed cleavage shortly after initiation, a trend we also observe in various tissues of mouse and human⁹ (Supplementary Fig. 2). Inspection of gene tracks at the *Pigt* locus reveals upstream antisense cleavage shortly after a PAS (AATAAA) less than 400 bases from the *Pigt* TSS, whereas in the sense direction cleavage is confined to the TES (Fig. 1b). Similar patterns were observed for subsets of promoters (promoters without nearby genes, global run-on sequencing (GRO-seq)-defined divergent promoters, and chromatin immunoprecipitation sequencing (ChIP-seq)-defined RNAPII phosphorylated at Ser 5-occupied promoters)¹⁰, or for high-confidence cleavage sites, cleavage reads and cleavage clusters (Supplementary Fig. 3). Of all divergent promoters, nearly half (48%) produce PAS-dependent upstream antisense cleavage events within 5 kb of the coding-gene TSS, compared to 33% downstream of the TSS. We validated several of these promoter-proximal sense and antisense cleavage sites using rapid amplification of 3' cDNA ends (3'-RACE) (Supplementary Fig. 4).

Similar to annotated cleavage sites at TESs of genes, these upstream antisense cleavage sites are associated with the PAS located at the expected position, about 22 nucleotides upstream of the cleavage site (Supplementary Fig. 5a, b)^{11,12}. Moreover, the nucleotide sequence composition flanking the cleavage sites resembles that of TESs of genes (Supplementary Fig. 5c–e), including a downstream U-rich region^{13,14}. To determine whether members of the canonical cleavage and polyadenylation machinery bind specifically to uaRNA cleavage sites, we analysed available crosslinking immunoprecipitation (CLIP) sequencing data sets for ten human canonical 3'-end-processing factors, namely CPSF160 (also known as CPSF1), CPSF100 (CPSF2), CPSF73 (CPSF3), CPSF30 (CPSF4), Fip1 (FIP1L1), CstF64 (CSTF2) and its paralogue CstF64 τ (CSTF2T), CF Im25 (NUDT21), CF Im59 (CPSF7) and CF Im68 (CPSF6), along with poly(A) 3'-end sequencing data generated in HEK293 cells¹⁵. We detected specific binding of all ten factors at uaRNA cleavage sites with positional profiles identical or very similar to that of mRNA cleavage sites (Supplementary Fig. 6). These results indicate that the poly(A) tails we analysed were products of PAS-dependent cleavage and polyadenylation, rather than either a priming artefact or a PAS-independent polyadenylation representing a transient signal for RNA degradation^{16–18}.

¹David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ²Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. ³Computational and Systems Biology Graduate Program, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

*These authors contributed equally to this work.

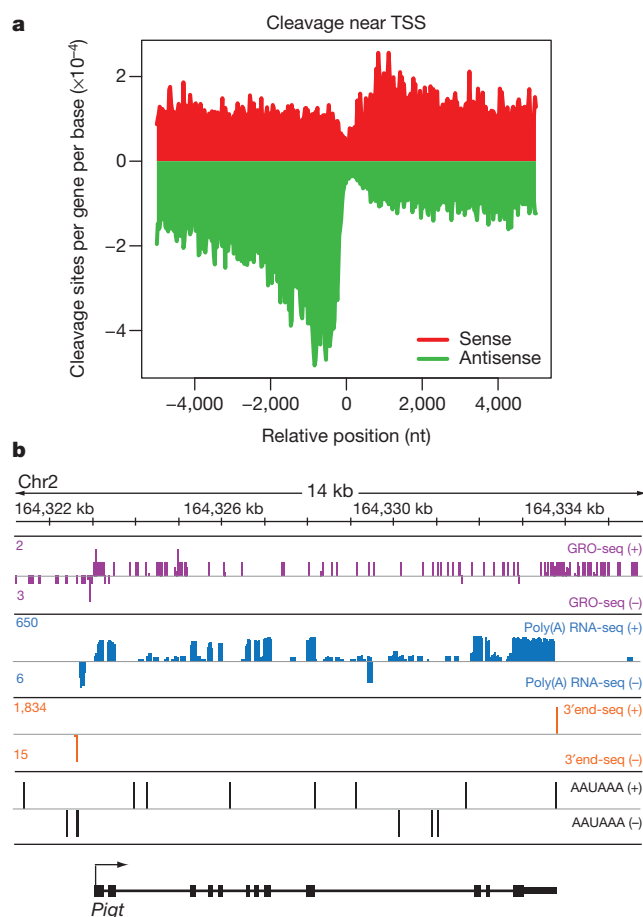


Figure 1 | Promoter-proximal PAS-dependent termination of uaRNA.

a, Metagenome plot of sense (red) or antisense (green) unique cleavage sites flanking coding-gene TSS. The number of unique cleavage sites per gene per base in each 25-base-pair (bp) bin across 5 kb upstream and downstream of the TSS in nucleotides (nt) is plotted. Mean cleavage density of first 2 kb: sense/antisense = 1.45/3.10. **b**, Genome browser view from the *Pigf* locus (shown in black on the + strand) displaying the following tracks with + strand (top) and - strand (bottom) represented: GRO-seq (purple)²⁸, Poly(A) + RNA-seq (blue)²⁹, 3' end RNA-seq (orange) and PAS (AAUAAA, black). For each gene track, the x axis represents the linear sequence of genomic DNA. The numbers on the left-hand side represent the maximum read density on each track.

As a first step to understand the molecular mechanism underlying the cleavage bias, we examined the frequency of PAS in a 6-kb region on the four strands flanking the coding-gene TSS. We observed an approximately 33% depletion of the canonical AATAAA PAS hexamer specifically downstream of the TSS on the coding strand of genes as compared to the other regions (Fig. 2a). As this 33% depletion is unlikely to explain the twofold cleavage bias observed (see simulation results in Supplementary Fig. 8a), we searched for additional discriminative hexamer sequence signals in an unbiased manner. All 4,096 hexamers were ranked by enrichment in the first 1 kb of the sense strand of genes relative to the corresponding upstream antisense region (Fig. 2b). Interestingly, we identified the PAS as the most depleted sequence in sense genes relative to the upstream antisense region of gene TSSs. In addition, we identified 5'-splice-site-related sequences (or sequences recognized by U1 snRNP, referred to as U1 sites) as the most enriched hexamers in sense genes (Fig. 2b) relative to antisense regions. This includes the consensus GGUAAG (first) that is perfectly complementary to the 5' end of the U1 snRNA, as well as GGUGAG (third) and GUGAGU (fifth), which represent common 5' splice site sequences (with the first GU in each motif located at the intron start). Consistent with the hexamer enrichment analysis, a metagenome plot displaying an unbiased prediction of strong, medium and weak U1 sites (see Methods)

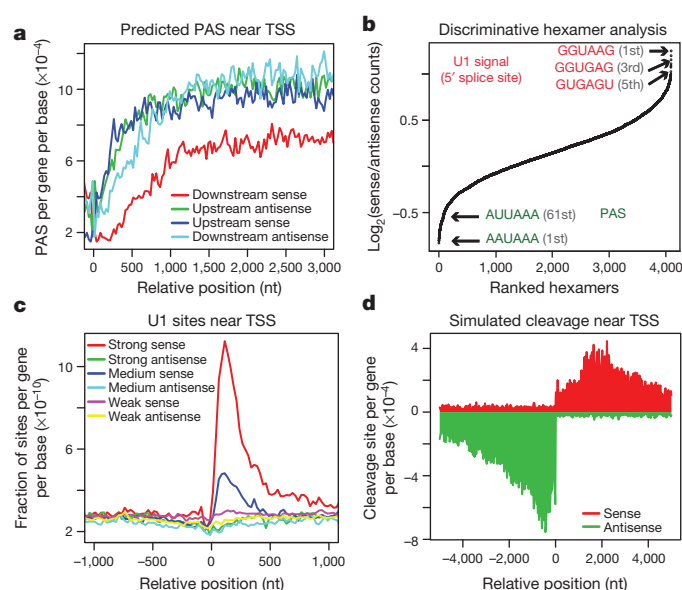


Figure 2 | Asymmetric distribution of PAS and U1 signals flanking coding-gene TSS. **a**, Number of AATAAA sites per gene per base in each 25-bp bin within a 3-kb region flanking the gene TSS on the downstream sense (red), downstream antisense (light blue), upstream antisense (green) and upstream sense (dark blue) strands. **b**, Rank of all 4,096 hexamers by enrichment (\log_2 ratio) in the first 1 kb of all coding genes in the sense direction relative to 1 kb in the upstream antisense direction of the TSS. **c**, Density of predicted 5' splice sites within a 1-kb region flanking gene TSS. Strong, medium and weak 5' splice sites are defined in Methods. **d**, Metagenome plot of simulated cleavage sites around gene TSS. The first unprotected PAS (AAUAAA) that is not within 1 kb downstream of a strong U1 site for all coding genes is plotted. Mean cleavage density of first 2 kb: sense/antisense = 2.08/4.99.

revealed strong enrichment of U1 signals in the first 500 base pairs downstream of the TSS, with essentially only background levels observed in all other regions and a small depletion in the upstream antisense direction (Fig. 2c).

The asymmetric distribution of U1 sites and PAS sites flanking the TSS could potentially explain the biased cleavage pattern shown in Fig. 1a if the U1 snRNP complex suppresses cleavage and polyadenylation near a U1 site, as has been observed in various species including human and mouse^{19–21}. Consistent with this model, we observed a depletion of cleavage sites, especially frequent cleavage sites, downstream of strong U1 sites (Supplementary Fig. 7a). Focusing on the upstream antisense direction, the presence of proximal PAS sites (within 1 kb of the coding-gene TSS) is significantly associated with shorter uaRNAs ($P < 10^{-15}$), whereas the presence of proximal U1 sites is significantly associated with longer uaRNAs but only in the presence of proximal PAS sites ($P < 0.0006$), consistent with a model in which the U1 snRNP promotes RNA lengthening by suppressing proximal PASs (Supplementary Fig. 7b). To test whether the encoded bias in U1 and PAS signal distribution explains the cleavage bias observed from our 3'-end sequencing analysis, we performed a cleavage site simulation using predicted strong U1 sites and canonical PAS (AATAAA) sequences. Specifically, we defined a protection zone of 1 kb downstream of a strong U1 site and used the first unprotected PAS as the cleavage site. The metagenome plot of simulated cleavage events (Fig. 2d) recapitulated the main features of the observed distribution (Fig. 1a), including an antisense peak around 700 bases upstream and a ~twofold difference between sense and antisense strands. Similar patterns were robustly observed when varying the size of the protection zone (Supplementary Fig. 8). Thus, we identified a U1–PAS axis flanking gene promoters that may explain why uaRNAs undergo early termination.

To validate the U1–PAS axis model, we functionally inhibited the U1 snRNP in mouse ESCs. Specifically, we transfected ESCs with either an antisense morpholino oligonucleotide (AMO) complementary to

the 5' end of the U1 snRNA to block its binding to 5' splice sites (or similar sequences) or a control AMO with scrambled sequences followed by 3'-end RNA sequencing^{19,20}. Interestingly, we observed in two biological replicates a marked increase in promoter-proximal cleavage events in coding genes but only a slight increase in upstream antisense regions, which eliminates the asymmetric bias in promoter-proximal cleavage we observed in either the wild-type cells or the cells treated with scrambled control AMOs (Fig. 3). These observations confirm that U1 snRNP protects sense RNA in protein-coding genes from premature cleavage and polyadenylation in promoter-proximal regions, thus reinforcing transcriptional directionality of genes. However, in the antisense direction, the activity of the U1 snRNP is much lower and there is little enhancement in cleavage sites upon U1 snRNP inhibition.

The conservation of the asymmetric cleavage pattern across human and mouse (Supplementary Fig. 2) led us to examine whether there is evolutionary selection on the U1–PAS axis. Previously, mouse protein-coding genes have been assigned to 12 evolutionary branches and dated by analysing the presence or absence of orthologues in the vertebrate phylogeny²². We find strong trends of progressive gain of U1 sites depending on the age of a gene (Fig. 4a) and loss of PAS sites (Fig. 4b) over time at the 5' end (the first 1 kb) of protein-coding genes, suggesting that suppression of promoter-proximal transcription termination is important for maintaining gene function. Interestingly, the same trends, although weaker, are observed in upstream antisense regions, suggesting that at least a subset of uaRNAs may be functionally important, in that over time they gain U1 sites and lose PAS sites to become more extensively transcribed. In addition to the coding strand of genes (downstream sense region), PAS sites were also progressively lost on the other three strands flanking TSS (Fig. 4b). This observation probably reflects the increases in CpG-rich sequences within 1 kb of gene TSSs and suggests that coding genes acquire CpG islands as they age (Fig. 4c). However, the bias of low-PAS-site density in the sense direction extends across the total transcription unit (Supplementary Fig. 9) and is distinct from the CpG density near the promoter.

We also propose that some long noncoding RNAs (lncRNAs) generated from bidirectional promoters might represent an evolutionary

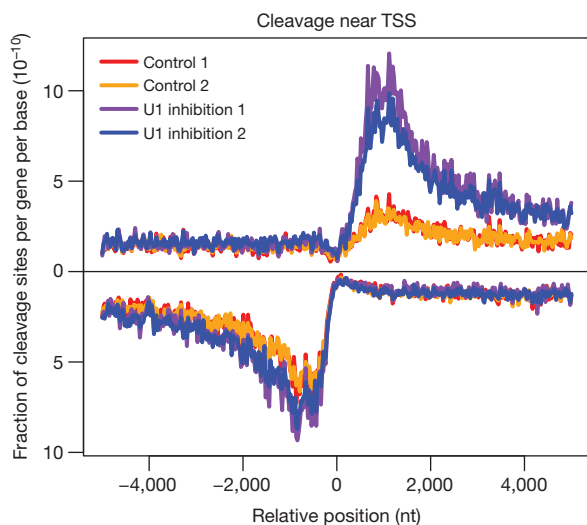


Figure 3 | Promoter-proximal cleavage sites are altered upon functional U1 inhibition. *y* axis represents the number of cleavage sites per gene per base divided by the total number of cleavage sites identified in each 3'-end sequencing library in a 5-kb region flanking the coding-gene TSS. Signal for the antisense strand is set as negative. U1 inhibition 1 (purple) and U1 inhibition 2 (blue) represent 3'-end sequencing libraries generated from mouse ESCs treated with a U1-targeting AMO. Control 1 (red) and control 2 (orange) represent 3'-end sequencing libraries generated from mouse ESCs treated with a scrambled control AMO. Mean cleavage density of first 2 kb: sense/antisense = 2.5/4.4 (control 1), 2.4/4.3 (control 2), 7.0/5.8 (U1 inhibition 1), 5.9/5.5 (U1 inhibition 2).

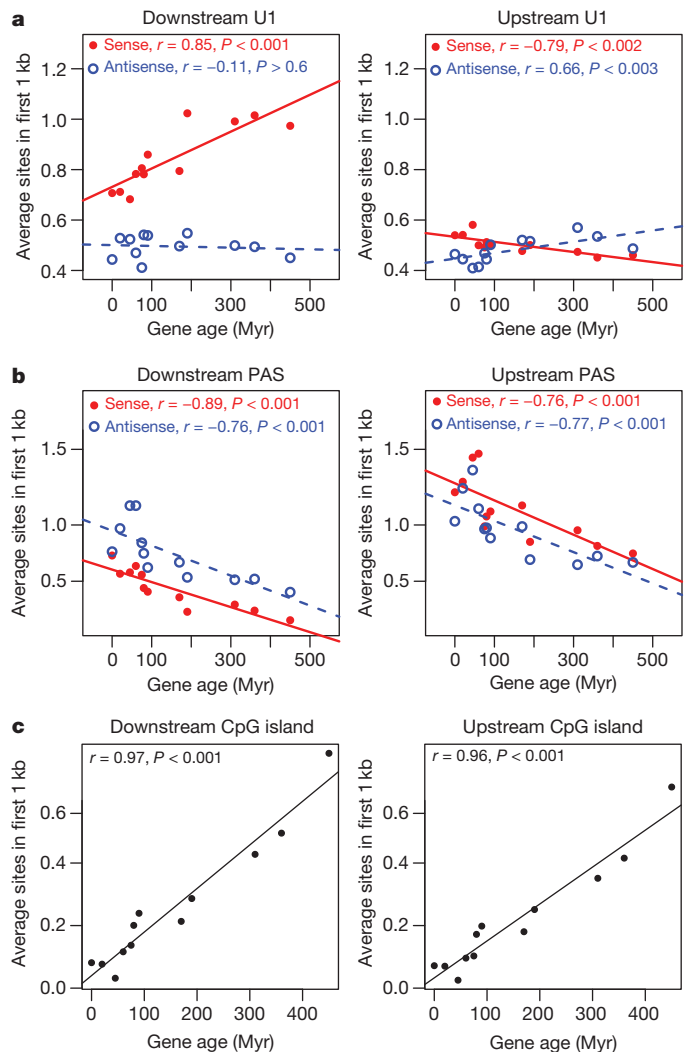


Figure 4 | Evolutionary gain and loss of U1 and PAS sites. **a**, Average number of strong U1 sites in the first 1 kb of protein-coding genes and upstream regions. **b**, Average number of PAS sites in the first 1 kb downstream and upstream of coding-gene TSS, respectively. **c**, Average number of CpG islands overlapping the first 1 kb of protein-coding genes and upstream regions. Genes are divided into 12 ordered groups by gene age. *x* axis indicates the age (Myr, million years) of gene groups. Number of genes in each group (from old to young): 11,934, 1,239, 914, 597, 876, 1,195, 279, 175, 198, 315, 926 and 1,143. Solid red dots and blue circles indicate sites on the sense and antisense strands, respectively. Pearson correlation coefficient (*r*) between gene age and average site counts is calculated and a line is fitted by linear regression. *P* values are estimated from 1,000 randomizations of gene age assignments.

intermediate between uaRNAs and protein-coding genes. Consistent with this, annotated head-to-head mRNA–lncRNA pairs as a whole showed a bias (in terms of promoter-proximal cleavage site, U1 site and PAS site distributions flanking the coding-gene TSS) weaker than head-to-head mRNA–uaRNA pairs but stronger than mRNA–mRNA pairs (Supplementary Fig. 10). This is also consistent with recent results suggesting that *de novo* protein-coding genes originate from lncRNAs at bidirectional promoters²³.

The U1–PAS axis probably has a broader role in limiting pervasive transcription throughout the genome. The enrichment of U1 sites and depletion of PAS sites are confined to the sense strand within the gene body, whereas intergenic and antisense regions show relatively high PAS but low U1 density (Supplementary Fig. 9), indicating that the U1–PAS axis may serve as a mechanism for terminating transcription in both antisense and intergenic regions.

Together, we propose that a U1–PAS axis is important in defining the directionality for transcription elongation at divergent promoters (Supplementary Fig. 11). Although the U1–PAS axis may explain the observed cleavage bias at promoters surprisingly well, it seems probable that additional *cis*-elements may influence PAS usage²⁴ and will need to be integrated into this model. There may also be other PAS-independent mechanisms that contribute to the termination of transcription in upstream antisense regions and across the genome^{25–27}. However, evidence for the U1–PAS axis is found in several different tissues of mouse and human, indicating its wide use as a general mechanism to regulate transcription elongation in mammals. Like protein-coding transcripts, lncRNAs must also contend with the U1–PAS axis. These RNAs and short non-coding RNAs from divergent transcription of gene promoters may be considered to be part of a continuum that varies in the degree of U1–PAS axis activity.

METHODS SUMMARY

Total RNA was extracted from V6.5 mouse ESCs that were grown under standard ESC culture conditions². Poly(A) RNA was selected, fragmented using a limited RNase T1 digestion and reverse transcribed using an oligo-dT-containing primer, and the resulting cDNA was circularized and PCR amplified using Illumina-specific primers. U1 inhibition experiments were performed as previously described^{19,20}.

Full Methods and any associated references are available in the online version of the paper.

Received 4 February; accepted 4 June 2013.

Published online 23 June 2013.

- Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
- Seila, A. C. *et al.* Divergent transcription from active promoters. *Science* **322**, 1849–1851 (2008).
- Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
- Preker, R. *et al.* RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851–1854 (2008).
- Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature Rev. Genet.* **13**, 720–731 (2012).
- Flynn, R. A., Almada, A. E., Zamudio, J. R. & Sharp, P. A. Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proc. Natl Acad. Sci. USA* **108**, 10460–10465 (2011).
- Preker, R. *et al.* PROMoter uPstream Transcripts share characteristics with mRNAs and are produced upstream of all three major types of mammalian promoters. *Nucleic Acids Res.* **39**, 7179–7193 (2011).
- Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770–1782 (2011).
- Derti, A. *et al.* A quantitative atlas of polyadenylation in five mammals. *Genome Res.* **22**, 1173–1183 (2012).
- Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
- Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M. & Gautheret, D. Patterns of variant polyadenylation signal usage in human genes. *Genome Res.* **10**, 1001–1010 (2000).
- Tian, B., Hu, J., Zhang, H. & Lutz, C. S. A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **33**, 201–212 (2005).
- Gil, A. & Proudfoot, N. J. Position-dependent sequence elements downstream of AAUAAA are required for efficient rabbit β -globin mRNA 3' end formation. *Cell* **49**, 399–406 (1987).
- MacDonald, C. C., Wilusz, J. & Shenk, T. The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Mol. Cell. Biol.* **14**, 6647–6654 (1994).
- Martin, G., Gruber, A. R., Keller, W. & Zavolan, M. Genome-wide analysis of pre-mRNA 3' end processing reveals a decisive role of human cleavage factor I in the regulation of 3' UTR length. *Cell Rep.* **1**, 753–763 (2012).
- LaCava, J. *et al.* RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121**, 713–724 (2005).
- Wyers, F. *et al.* Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121**, 725–737 (2005).
- Vaňáčová, S. *et al.* A new yeast poly(A) polymerase complex involved in RNA quality control. *PLoS Biol.* **3**, e189 (2005).
- Berg, M. G. *et al.* U1 snRNP determines mRNA length and regulates isoform expression. *Cell* **150**, 53–64 (2012).
- Kaida, D. *et al.* U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668 (2010).
- Andersen, P. K., Lykke-Andersen, S. & Jensen, T. H. Promoter-proximal polyadenylation sites reduce transcription activity. *Genes Dev.* **26**, 2169–2179 (2012).
- Zhang, Y. E., Vrbancin, M. D., Landback, P., Marais, G. A. & Long, M. Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X chromosome. *PLoS Biol.* **8**, e1000494 (2010).
- Xie, C. *et al.* Hominoid-specific *de novo* protein-coding genes originating from long non-coding RNAs. *PLoS Genet.* **8**, e1002942 (2012).
- Hu, J., Lutz, C. S., Wilusz, J. & Tian, B. Bioinformatic identification of candidate *cis*-regulatory elements involved in human mRNA polyadenylation. *RNA* **11**, 1485–1493 (2005).
- Connelly, S. & Manley, J. L. A CCAAT box sequence in the adenovirus major late promoter functions as part of an RNA polymerase II termination signal. *Cell* **57**, 561–571 (1989).
- Arigo, J. T., Eyler, D. E., Carroll, K. L. & Corden, J. L. Termination of cryptic unstable transcripts is directed by yeast RNA-binding proteins Nrd1 and Nab3. *Mol. Cell* **23**, 841–851 (2006).
- Zhang, L., Ding, Q., Wang, P. & Wang, Z. An upstream promoter element blocks the reverse transcription of the mouse insulin-degrading enzyme gene. *Biochem. Biophys. Res. Commun.* **430**, 26–31 (2013).
- Min, I. M. *et al.* Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* **25**, 742–754 (2011).
- Sigova, A. A. *et al.* Divergent transcription of lncRNA/mRNA gene pairs in embryonic stem cells. *Proc. Natl Acad. Sci. USA* **110**, 2876–2881 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors acknowledge the service to the MIT community of the late Sean Collier. We would like to thank N. Spies for generously sharing his optimized 3'-end sequencing protocol, C. Lin for providing computational assistance, M. Lindstrom for assistance on constructing Supplementary Figure 11, and S. Chen, A. Chiu, M. Jangi, Q. Liu, J. Wilusz and J. Zamudio for reading of the manuscript. We also thank the Core Facility in the Swanson Biotechnology Center at the David H. Koch Institute for Integrative Cancer Research at MIT for their assistance with high-throughput sequencing. This work was supported by United States Public Health Service grants R01-GM34277 and R01-CA133404 from the National Institutes of Health (P.A.S.), partially by Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute, and by a Public Health Service research grant (GM-085319) from the National Institute of General Medical Sciences (C.B.B.). X.W. is a Howard Hughes Medical Institute International Student Research fellow.

Author Contributions A.E.A., X.W. and P.A.S. conceived and designed the research. A.E.A. performed experiments. X.W. and A.J.K. performed computational analysis. A.E.A., X.W., C.B.B. and P.A.S. analysed the data and wrote the manuscript.

Author Information 3'-end sequencing data is deposited in the Gene Expression Omnibus under accession number GSE46433. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.A.S. (sharppa@mit.edu).

METHODS

Cell culture. V6.5 (C57BL/6-129) mouse ESCs (Koch Institute Transgenic Facility) were grown under standard ESC culture conditions².

Poly(A) 3'-end sequencing. Total RNA was extracted from V6.5 mouse ESCs using Ambion's Ribopure kit. Poly(A)-selected RNA was fragmented using Invitrogen's RNase T1 (biochemistry grade). Reverse transcription was performed with an RT oligo (Supplementary Table 1) at 0.25 μ M final concentration using Invitrogen's Superscript III Reverse Transcriptase according to the manufacturer's protocol. The resulting cDNA was run on a 6% TBE-Urea polyacrylamide gel (National Diagnostics) and the 100–300 size range of products were gel extracted and eluted overnight. The gel-purified cDNA products were circularized using Epicentre's CircLigase II according to the manufacturer's protocol. Circularized cDNA was PCR amplified using the New England Biolab's Phusion High-Fidelity DNA Polymerase for 15–18 cycles using the primers described in Supplementary Table 1. Amplified products were run on a 1.5% agarose gel and the 200–400 size range was extracted using Qiagen's MinElute Gel Extraction Kit. The 3'-end library was then submitted for Illumina sequencing on the Hi-Seq 2000 platform.

U1 inhibition with AMO. V6.5 mouse ESCs were transfected using the Amaxa Nucleofector II with program A-23 (mouse ESC specific) according to the manufacturer's protocol. Specifically, 2.5 million V6.5 ESCs were transfected with 7.5 μ M of U1-targeting or a scrambled AMO for 8 h^{19,20}, prior to RNA-sequencing analysis.

3'-RACE. Total RNA was extracted using Ambion's Ribopure kit and DNase-treated using Ambion's DNA Free-Turbo. 3'-RACE was performed using Ambion's Gene Racer Kit according to the manufacturer's instructions. 3'-end PCR products were run on a 1.5% agarose gel, gel extracted using Qiagen's gel extraction kit, and Sanger sequenced. All primers are described in Supplementary Table 1.

Reads mapping. Raw reads were processed with the program cutadapt²⁰ to trim the adaptor sequence (TGGAATTCTCGGGTCCCAAGGAAGTCCAGTCACA TCAC) from the 3' end. Reads longer than 15 nt after adaptor trimming were mapped to the mouse genome (mm9) with Bowtie³¹ requiring unique mapping with at most two mismatches (options: -n 2 -m 1 -best -strata). Mapped reads were collapsed by unique 3'-end positions.

Internal priming filter. To remove reads whose A-tail is encoded in the genome rather than added post-transcriptionally, we filtered reads that had: (1) more than ten As in the first 20-nt window or (2) more than six As in the first 10-nt window downstream from the detected cleavage site of the 3' end. The threshold used is based on the bimodal distribution of the number of As downstream of the annotated TES.

PAS filter. In addition to a set of 12 hexamers identified previously in mouse and human expressed sequence tag analysis^{11,12}, we analysed the annotated TES in the mouse genome to identify additional potential PAS variants. All hexamers with at most two mismatches to the canonical AATAAA motif were used to search in the sequence up to 100 nt upstream of the annotated TES. The distribution of the position of each hexamer relative to the TES (a histogram) is compared to that of AATAAA. Hexamers with a position profile similar to AATAAA will have a peak around position 20–24. We quantified the similarity using the Pearson correlation coefficient and used a cutoff of 0.5 after manual inspection. In total, 24 new hexamers were identified as potential PASs and a hierarchy was assigned for the 36 hexamers (PAS36): first, the 12 known variants were ranked by their frequency of usage in the mouse genome, and then the newly identified PAS ranked by their correlation with AATAAA in terms of the positional profile defined above. To define a window in which most PASs or variants are located, we searched for each of the 36 PAS variants within 100 nt of annotated gene 3' ends and chose the best one according to the designated hierarchy. We summarized the distance of the best PAS to the annotated TES and defined a window of 0–41 around the position 22 peak, such that 80% of the annotated TES have their best-matched PAS within that window. Using these criteria, we searched for PAS36 variants within the 0–41 window upstream of our experimentally sequenced 3' ends. If there were multiple PAS hexamers identified within this window for a given 3' end, we chose the best one defined by the hierarchy described above. Reads without any of the 36 PAS variants within the 0–41 window were discarded.

Remove potential false-positive cleavage sites. Owing to sequencing error, abundant transcripts such as ribosomal gene mRNAs can produce error-containing 3'-end reads that mapped to other locations in the genome, leading to false-positive cleavage sites. To remove such potential false-positive sites, we defined a set of 71,674 (7.5%) abundant cleavage sites that are supported with more than 100 reads from the pooled library. A Bowtie reference index was built using sequences within 50 nt upstream of those abundant sites. Nonabundant sites within these 50-nt reference regions were not used to search for false positives. Reads initially mapped to sites outside of these reference regions were re-mapped against the new index allowing up to two mismatches. Reads mapped to any of the reference regions in this analysis were treated as potential false-positive reads. Cleavage sites containing only potential false-positive reads were defined as potential false-positive sites and

were removed from subsequent analysis. In total, 7.2% (389,185) of initially mapped reads were outside of the reference regions. 0.34% of all mapped reads were classified as potential false-positive reads and 9.1% (86,425) of all cleavage sites were identified as potential false-positive sites.

Remove mouse B2 short interspersed elements (SINE) RNA-associated cleavage sites. We further removed cleavage sites associated with B2_Mm1a and B2_Mm1t SINE RNAs. These B2 SINE RNAs are transcribed by RNA Pol III but contain AAUAAA sequences near the 3' end. In total, 3.5% (33,696) of all cleavage sites passing the internal priming filter and the PAS filter were mapped within B2 regions or within 100 nt downstream of B2 3' end. These sites were removed.

Prediction of U1 sites/putative 5' splice sites. A nucleotide frequency matrix of 5' splice sites (3 nt in exon and 6 nt in intron) was compiled using all annotated constitutive 5' splice sites in the mouse genome. The motif was then used by FIMO³² to search significant matches ($P < 0.05$) on both strands of the genome. Matches were then scored by a maximum entropy model³³. Maximum entropy scores for all annotated 5' splice sites were also calculated to define thresholds used to classify the predicted sites into strong, medium and weak. Sites with scores larger than the median of annotated 5' splice sites (8.77) were classified as 'strong'. Sites with scores lower than 8.77 but higher than the threshold dividing the first and second quarter of annotated 5' splice sites (7.39) were classified as 'medium', and the rest of the predicted sites with scores higher than 4 were classified as 'weak'. Sites with scores lower than 4 were discarded.

Define a set of divergent promoters. GRO-seq data from mouse ESCs²⁸ were used to define a set of active and divergent promoters. Active promoters were defined as promoters with GRO-seq signal detected within the first 1 kb downstream of the sense strand. A promoter was considered divergent if it contained GRO-seq signal in the first 1 kb downstream of the sense strand and within the first 2 kb of the upstream antisense strand. A minimum number of two reads within the defined window (downstream 1 kb or upstream 2 kb) was used as a cutoff for background signals.

Define Ser 5-phosphorylated (Ser5P) RNAPII-bound TSS. ChIP-seq data for Ser5P RNA Pol II and corresponding input was downloaded from the Gene Expression Omnibus (GEO) database (accession number GSE20530 (ref. 10)) and peaks called using MACS³⁴ with default settings. TSSs less than 500 bp away from a peak summit were defined as bound.

Discriminative hexamer analysis. An unbiased exhaustive enumeration of all 4,096 hexamers was performed to find hexamers that are discriminative of downstream sense and upstream antisense strands of protein-coding gene promoters. Specifically, the first 1,000 nucleotides downstream sense and upstream antisense of all protein-coding gene TSSs were extracted from repeat masked genome (from UCSC genome browser, non-masked genome sequence gave similar results). For each hexamer, the total number of occurrences on each side was counted and the log₂ ratio of the occurrences on sense versus antisense strand was calculated as a measure of enrichment on the sense but depletion on the antisense strand.

Cleavage site simulation. Protein-coding genes and 10-kb upstream antisense regions were scanned for strong U1 sites and PAS sites (AATAAA). Starting from protein-coding gene TSSs, the first unprotected PAS was predicted to be the cleavage site. A PAS is protected only if it is within a designated protection window (in nucleotides) downstream (+) of a strong U1 site.

Binding of 3'-end-processing factors in uRNA regions. RNA 3'-end cleavage and polyadenylation sites and CLIP-seq read density of ten 3'-end-processing factors in wild-type HEK293 cells were downloaded from GEO data set GSE37401. A cleavage site is defined as a uRNA cleavage site if it is outside any protein-coding gene but locates within 5 kb upstream antisense of a protein-coding gene. mRNA cleavage sites are defined as cleavage sites within 100 bases of annotated protein-coding gene ends. For each 3'-end-processing factor, CLIP read density within 200 bases of all cleavage sites are added up every 5-bp bin and then normalized such that the max value is 1.

Evolutionary analysis of U1 sites, PAS sites and CpG islands. Mouse protein-coding gene branch/age assignment was obtained from a previous analysis²². The number of strong U1 sites, PAS (AATAAA) sites and CpG islands (UCSC mm9 annotations) in the first 1-kb region flanking the TSS on each strand were calculated, and the average number of sites in each branch/age group was plotted against gene age. The Pearson correlation coefficient and linear regression fitting were done using R. Significance of the correlation was assessed by comparing to a null distribution of correlation coefficients calculated by shuffling gene branch/age assignments 1,000 times.

Bidirectional promoter analysis. For each annotated TSS the closest upstream antisense TSS was identified and those TSS pairs within 1 kb were defined as head-to-head pairs. LncRNAs were defined as noncoding RNAs longer than 200 bp. UCSC mm9 gene annotations were used in this analysis.

30. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).

31. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
32. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017–1018 (2011).
33. Yeo, G. & Burge, C. B. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* **11**, 377–394 (2004).
34. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

Ultrasensitive fluorescent proteins for imaging neuronal activity

Tsai-Wen Chen¹, Trevor J. Wardill^{1†}, Yi Sun¹, Stefan R. Pulver¹, Sabine L. Renninger², Amy Baohan^{1,3}, Eric R. Schreiter¹, Rex A. Kerr¹, Michael B. Orger², Vivek Jayaraman¹, Loren L. Looger¹, Karel Svoboda¹ & Douglas S. Kim¹

Fluorescent calcium sensors are widely used to image neural activity. Using structure-based mutagenesis and neuron-based screening, we developed a family of ultrasensitive protein calcium sensors (GCaMP6) that outperformed other sensors in cultured neurons and in zebrafish, flies and mice *in vivo*. In layer 2/3 pyramidal neurons of the mouse visual cortex, GCaMP6 reliably detected single action potentials in neuronal somata and orientation-tuned synaptic calcium transients in individual dendritic spines. The orientation tuning of structurally persistent spines was largely stable over timescales of weeks. Orientation tuning averaged across spine populations predicted the tuning of their parent cell. Although the somata of GABAergic neurons showed little orientation tuning, their dendrites included highly tuned dendritic segments (5–40- μ m long). GCaMP6 sensors thus provide new windows into the organization and dynamics of neural circuits over multiple spatial and temporal scales.

Neural activity causes rapid changes in intracellular free calcium^{1–4}. Calcium imaging experiments have relied on this principle to track the activity of neuronal populations^{5,6} and to probe excitation of small neurons and neuronal microcompartments^{2,7–10}. Genetically encoded protein sensors can be targeted to specific cell types^{2,9,11,12} for non-invasive imaging of identified neurons and neuronal compartments^{8,13–15} over chronic timescales⁶.

Calcium indicator proteins include the single fluorophore sensor GCaMP (refs 11, 16, 17) and several families of Förster resonance energy transfer based sensors^{18–22}. However, none of these protein-based indicators have yet surpassed the sensitivity and speed of commonly

used synthetic calcium indicators (for example, Oregon Green Bapta-1-AM, OGB1-AM). Therefore, depending on the experimental goals, investigators choose between sensitive synthetic indicators delivered by invasive chemical or physical methods, or less sensitive protein sensors delivered by genetic methods.

Multiple rounds of structure-guided design have made GCaMPs the most widely used protein calcium sensors^{11,16,17}. But past efforts in optimizing GCaMPs and other indicators of neuronal function were limited by the throughput of quantitative and physiologically relevant assays. Because neurons have unusually fast calcium dynamics and low peak calcium accumulations⁴, sensors designed to probe neuronal

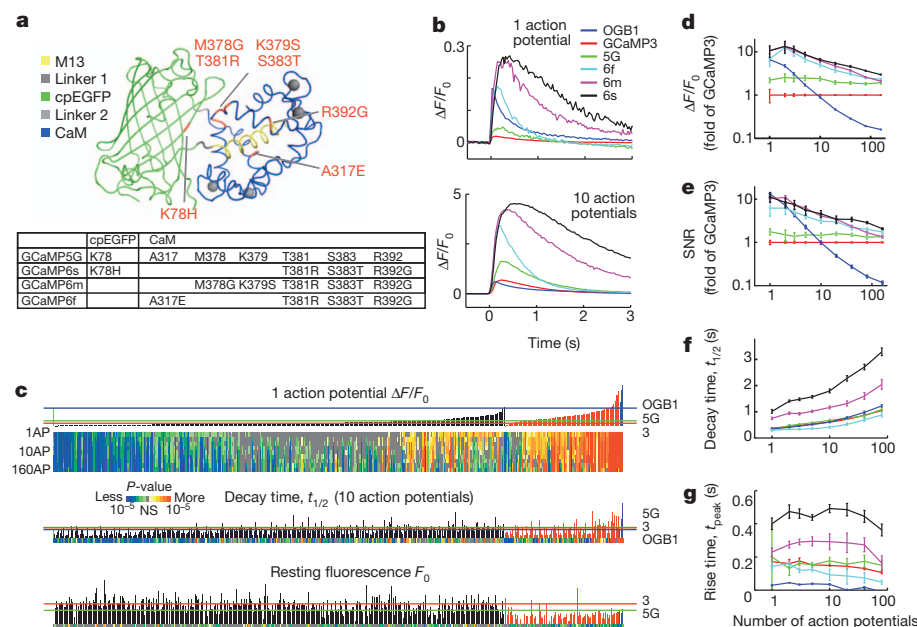


Figure 1 | GCaMP mutagenesis and screening in dissociated neurons. **a**, GCaMP structure^{27,51} and mutations in different GCaMP variants relative to GCaMP5G. **b**, Responses averaged across multiple neurons and wells for GCaMP3, 5G, 6f, 6m, 6s, and OGB1-AM. Top, fluorescence changes in response to 1 action potential. Bottom, 10 action potentials. **c**, Screening results, 447 GCaMPs. Top, fluorescence change in response to 1 action potential (vertical bars, $\Delta F/F_0$; green bar, OGB1-AM, left; black bars, single GCaMP mutations; red bars, combinatorial mutations; blue, GCaMP6 indicators) and significance values for different action potential stimuli (colour plot). Middle, half decay time after 10 action potentials. Bottom, resting fluorescence, F_0 normalized to nuclear mCherry fluorescence. Red line, GCaMP3 level; green line, GCaMP5G level; blue line, OGB1-AM level. AP, action potential. **d–g**, Comparison of GCaMP sensors and OGB1-AM as a function of stimulus strength (colours as in **b**). **d**, Response amplitude. **e**, Signal-to-noise ratio (SNR). **f**, Half decay time. **g**, Time to peak (after stimulus offset). Error bars correspond to s.e.m ($n = 300, 16, 8, 11, 13$ and 11 wells for GCaMP3, GCaMP5G, OGB1-AM, 6f, 6m and 6s, respectively).

¹Janelia Farm Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, Virginia 20147, USA. ²Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, Avenida Brasília, Doca de Pedrouços, 1400-038, Lisbon, Portugal. ³Department of Neurobiology, University of California Los Angeles, Los Angeles, California 90095, USA. [†]Present address: Marine Biological Laboratory, Program in Sensory Physiology and Behavior, 7 MBL Street, Woods Hole, Massachusetts 02543, USA.

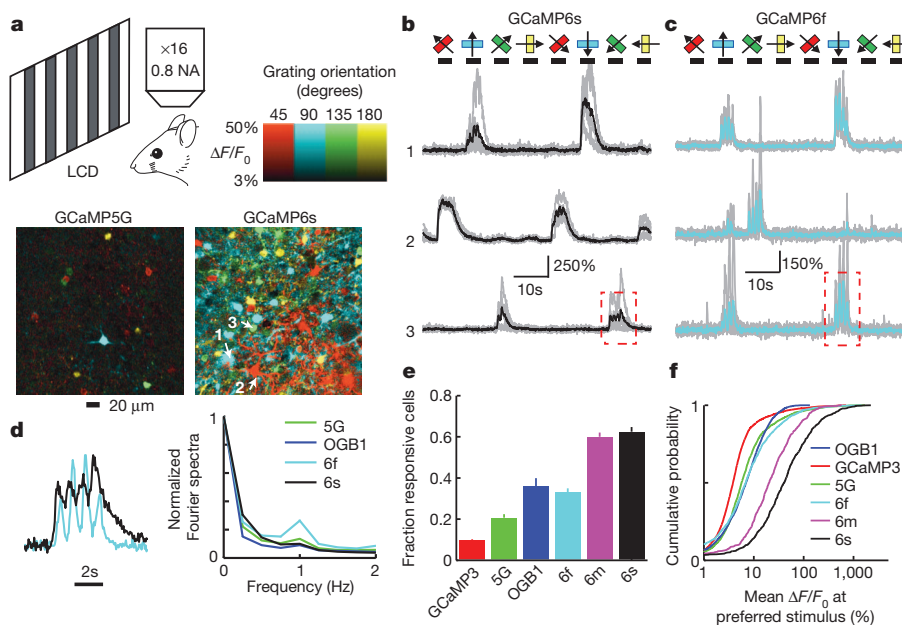


Figure 2 | GCaMP6 performance in the mouse visual cortex. **a**, Top, schematic of the experiment. Bottom, field of view showing neurons colour-coded according to their preferred orientation (hue) and response amplitude (brightness) for GCaMP5G (left) and GCaMP6s (right). **b**, Example traces from three neurons expressing GCaMP6s. Single sweeps (grey) and averages of 5 sweeps (black) are overlaid. Directions of grating motion (8 directions) are shown above traces (arrows). **c**, Example traces from three neurons expressing GCaMP6f. Single sweeps (grey) and averages of 5 sweeps (cyan) are overlaid. **d**, Left, high magnification view of fluorescence changes corresponding to the red boxes in **b** (black) and **c** (cyan), normalized to the peak of the response. Right, Fourier spectra normalized to the response amplitude at 0 Hz for neurons driven with 1 Hz drifting gratings, transduced with GCaMP5G, OGB1-AM, 6f or 6s. **e**, The fraction of cells scored as responding to visual stimulation when loaded with different calcium indicators. Error bars correspond to s.e.m. ($n = 70, 39, 23, 38, 21$ and 34 fields of view (FOVs) for GCaMP3, 5G, OGB1-AM, 6f, 6m and 6s, respectively). GCaMP3, 5G, and OGB1-AM data are from ref. 16. **f**, The distribution of fluorescence changes across cells at the preferred orientation.

function are best tested in neurons^{11,13,23,24}, rather than in non-neuronal systems, most of which show much slower and larger calcium changes¹⁹. We thus screened GCaMP variants produced by mutagenesis in neurons, and subsequently validated lead sensors in several *in vivo* systems.

GCaMP protein engineering

GCaMP (ref. 17) and its progeny^{11,16} consist of circularly permuted green fluorescent protein (cpGFP)²⁵, the calcium-binding protein calmodulin (CaM) and CaM-interacting M13 peptide²⁶ (Fig. 1a). The CaM-M13 complex is in proximity to the chromophore inside the cpGFP β -barrel²⁷. Calcium-dependent conformational changes in CaM-M13, including modulation of solvent access and the pK_a of the chromophore, cause increased brightness with calcium binding.

Despite extensive structure-guided optimization^{11,16}, GCaMP and other protein sensors still suffer from low sensitivity and slow kinetics.

We produced numerous additional GCaMP variants and tested them in automated neuronal assays (Fig. 1). With the aim of improving sensitivity, we focused mutagenesis on the interface between cpGFP and CaM at 16 amino acid positions, some mutagenized to near completion (Fig. 1a, Supplementary Table 5)¹⁶. Mutations were made at 18 additional sites, notably at the M13-CaM interface which can affect calcium affinity²⁸ (A317) and in CaM (R392) (ref. 16) (Fig. 1a).

Dissociated rat hippocampal neurons in 24-well plates were transduced with GCaMP variants (one per well), together with nuclear mCherry²⁹, using lentivirus-mediated gene transfer. Electrodes triggered trains of action potentials in all neurons within each well (Methods).

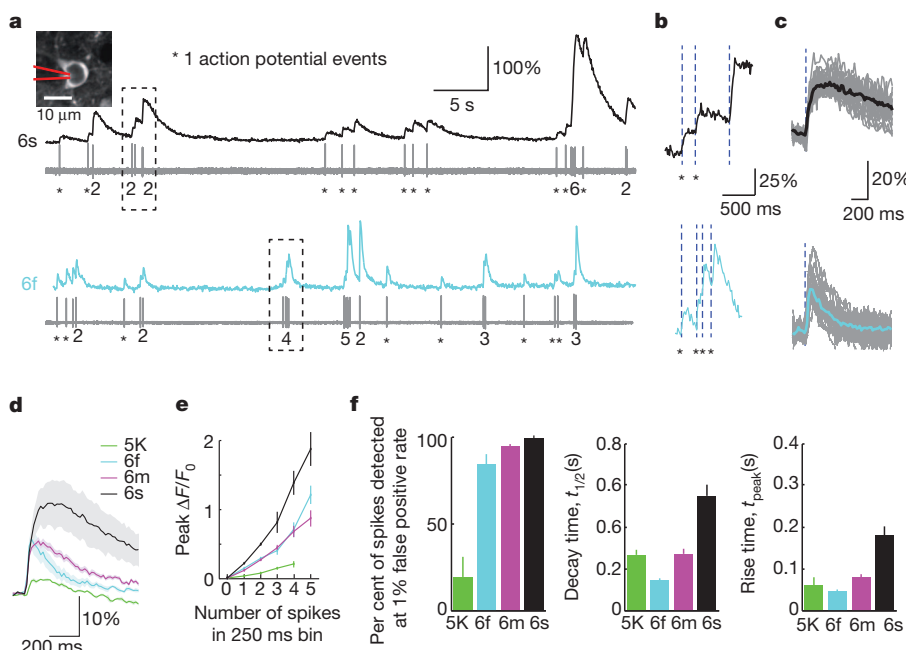


Figure 3 | Combined imaging and electrophysiology in the visual cortex.

a, Simultaneous fluorescence dynamics and spikes in a GCaMP6s (top) and a GCaMP6f (bottom) expressing neuron. The number of spikes for each burst is indicated below the trace (single spikes are indicated by asterisks). Left inset, a GCaMP6s-expressing neuron with the recording pipette indicated schematically. **b**, Zoomed-in view of bursts of action potentials. Top, GCaMP6s; bottom, GCaMP6f. **c**, Fluorescence change in response to one action potential. Top, GCaMP6s; bottom, GCaMP6f. **d**, Median fluorescence change in response to one action potential for different calcium indicators. Shading corresponds to s.e.m., $n = 9$ (GCaMP5K, data from ref. 16), 11 (GCaMP6f), 10 (GCaMP6m), 9 (GCaMP6s) cells. GCaMP5K and GCaMP5G have similar properties¹⁶. **e**, Peak fluorescence change as a function of number of action potentials in a 250 ms bin (5K: $n = 161, 65, 22, 4$ events for 1, 2, 3, 4 action potentials; 6f: $n = 366, 120, 50, 15, 7$ events for 1, 2, 3, 4, 5 action potentials; 6m: $n = 354, 105, 31, 11, 7$ events for 1, 2, 3, 4, 5 action potentials; 6s: $n = 250, 60, 20, 5, 4$ events for 1, 2, 3, 4, 5 action potentials). Error bars correspond to s.e.m. **f**, Comparison of GCaMP indicators. Left, fraction of isolated spikes detected at 1% false positive rate. Middle, half decay time. Right, rise time to peak. Error bars correspond to s.e.m.

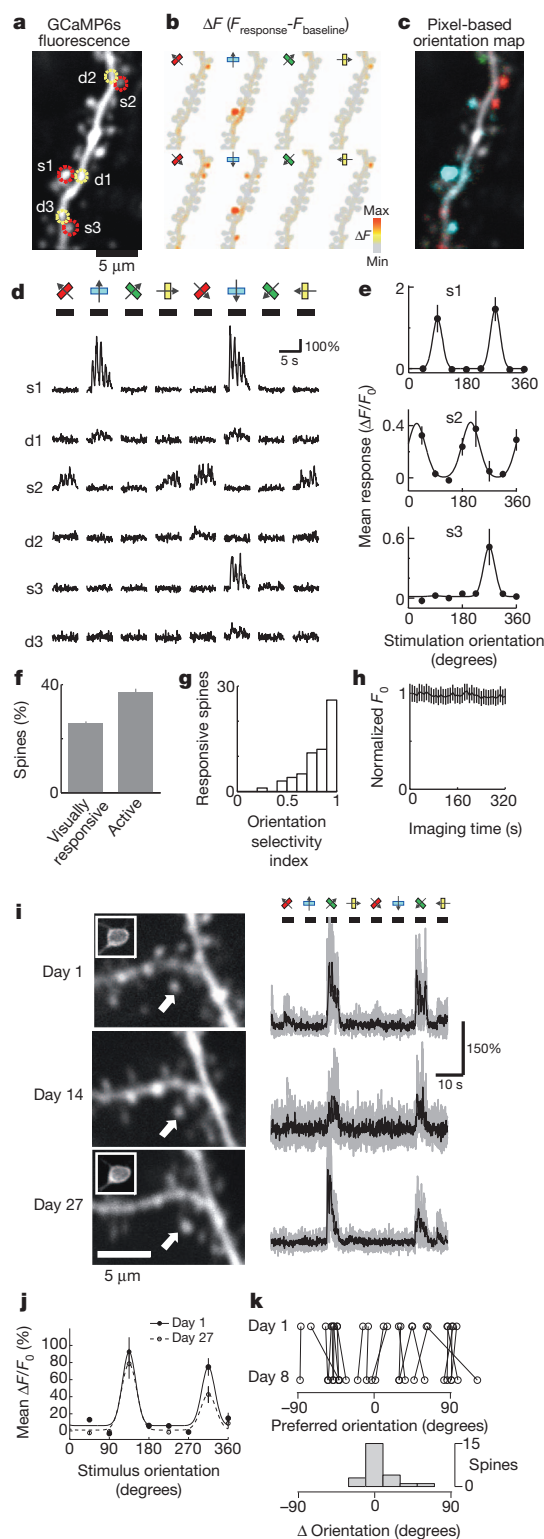


Figure 4 | Imaging activity in dendritic spines in the visual cortex. **a**, Image of an L2/3 dendritic branch expressing GCaMP6s. Regions of interest (ROIs) are indicated as dashed circles (red, spines; yellow, dendrites). **b**, Map of fluorescence change ($\Delta F = F_{\text{response}} - F_{\text{baseline}}$) in response to drifting gratings of 8 different orientations. **c**, Pixel-based map of orientation preference. **d**, Responses of dendritic spines (s1–s3) and neighbouring dendritic shafts (d1–d3) to drifting gratings with different orientations (corresponding to ROIs indicated in **a**). **e**, Orientation tuning of individual spines (s1, s2, s3). Error bars correspond to s.e.m. ($n = 5$ trials). **f**, Fraction of spines that show detectable calcium transients (active) and respond to visual stimulation (responsive) (see Methods for definitions) (228 spines; 15 dendrites; 4 mice). **g**, Distribution of the orientation selectivity index across visually responsive spines (62 spines). **h**, Baseline fluorescence across individual dendritic spines over 320 s of continuous imaging (228 spines; 15 dendrites; 4 mice; error bars reflect s.e.m. across spines). **i**, Left, the same GCaMP6s labelled spine imaged over weeks. Right, fluorescence responses to oriented drifting gratings. Insets, parent soma of imaged spines. **j**, Orientation selectivity of single spines measured over time (same as **i**). **k**, Top, preferred orientation for spines that responded in two imaging sessions separated by one week. Opposing stimulus directions are considered as equivalent in this analysis. Bottom, the distribution of $\Delta\text{Orientation}$ (difference in preferred orientation between two sessions).

addition, one mutation at the M13–CaM interface (A317E) accelerated kinetics (fourfold), but also reduced response amplitude (two-fold), compared to parent GCaMP variants (Supplementary Fig. 2). In a second round of mutagenesis, we combined beneficial mutations, selected based on improved response amplitudes (1–3 action potentials) and/or kinetics, without compromising baseline fluorescence or maximal response (160 action potentials) (94 variants, up to 8 beneficial point mutations; Fig. 1a, c, Supplementary Table 5). In some cases the beneficial effects were additive (Supplementary Fig. 2). A317E consistently accelerated the kinetics compared to the parent sensors. In total, we screened 447 GCaMP variants (Supplementary Table 5).

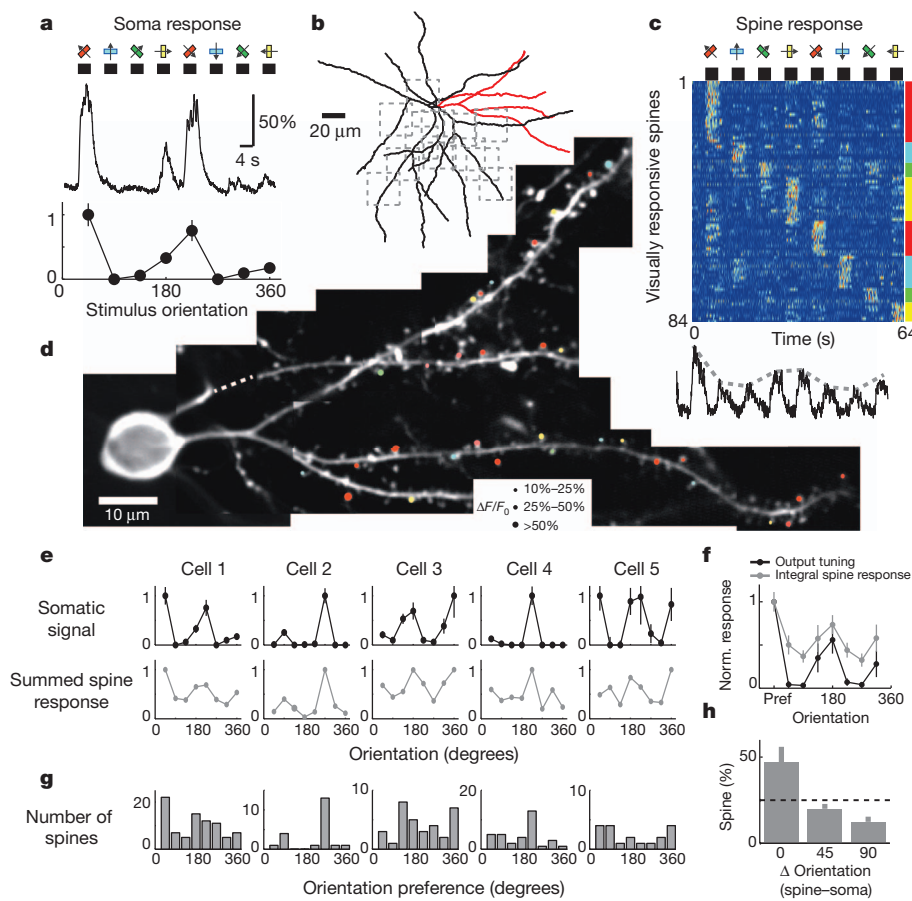
Based on screening in cultured neurons (Fig. 1), we chose three ultrasensitive GCaMP6 sensors (GCaMP6s, 6m, 6f; for slow, medium and fast kinetics, respectively) for characterization *in vivo*. These sensors vary in kinetics, with the more sensitive sensors having slower kinetics. Compared to GCaMP5G, the GCaMP6 sensors have similar baseline brightness and a 1.1- to 1.6-fold increase in dynamic range ($\Delta F/F_0$ at 160 action potentials). For small numbers of action potentials the most sensitive sensor, GCaMP6s, produced sevenfold larger signals (>10-fold larger than GCaMP3, Fig. 1b–e; Supplementary Table 1). Underlying this sensitivity gain are multiple factors (Supplementary Table 2). Compared to GCaMP5G, GCaMP6s exhibited threefold higher apparent affinity for calcium and 1.3-fold higher saturated fluorescence, with similar baseline fluorescence. Calcium-saturated GCaMP6s is 27% brighter than enhanced GFP (EGFP), its parent fluorescent protein. The fastest sensor, GCaMP6f, had twofold faster rise time and 1.7-fold faster decay time than GCaMP5G (Fig. 1f, g) (Supplementary Table 1). GCaMP6f is the fastest genetically-encoded calcium indicator for cytoplasmic free calcium in neurons, with sensitivity comparable to OGB1-AM (Fig. 1d–g). GCaMP6 indicators were more sensitive and/or faster than other GCaMP variants in presynaptic boutons of the *Drosophila* larval neuromuscular junction²⁴ (Supplementary Fig. 3; Supplementary Table 4), in projection neurons of the *Drosophila* adult antennal lobe during odour presentation¹⁶ (Supplementary Fig. 4), and in neuropil and somata of the zebrafish tectum *in vivo*¹⁶ (Supplementary Fig. 5).

Imaging neuronal populations in mouse V1

We next tested GCaMP6 in layer (L) 2/3 pyramidal neurons in the mouse visual cortex V1 *in vivo* (Fig. 2a). The majority of V1 neurons can be driven to fire action potentials in response to drifting gratings³⁰. V1 was infected with adeno-associated virus (AAV) expressing GCaMP variants (AAV-*hSyn1*-GCaMP variant)¹¹. Three weeks after AAV infection, the vast majority of L2/3 neurons showed fluorescence mainly in the neuronal cytoplasm (Supplementary Fig. 6). Sensory stimuli consisted of moving gratings presented in eight directions to the contralateral

Time-lapse images (35 Hz) of $\sim 800 \mu\text{m}$ fields of view containing 10–100 neurons were acquired, while delivering a series of action potential trains (Fig. 1b, c). Fluorescence changes extracted from single neurons were used to compare the sensitivity, dynamic range and kinetics of individual GCaMP variants and OGB1-AM (Fig. 1b–g). We monitored the resting brightness of the sensor by measuring green fluorescence relative to red mCherry fluorescence.

The assay revealed 43/348 point mutants with improved sensitivity compared to GCaMP3 ($\Delta F/F_0$ in response to one action potential; $P < 0.01$; Wilcoxon rank sum test) (Fig. 1c; Supplementary Fig. 1). In



eye^{12,16}. Two-photon imaging revealed visual stimulus-evoked fluorescence transients in subsets of neurons (Fig. 2a–c). These responses were stable across trials (Supplementary Fig. 8) and tuned to stimulus orientation (Fig. 2a, b and Supplementary Fig. 9). Orientation tuning was similar for GCaMP5G, GCaMP6f, GCaMP6m and bulk-loaded OGB1-AM (ref. 5) (Supplementary Fig. 9). Fluorescence transients were faster with GCaMP6f compared to other sensors and faithfully tracked dynamic sensory stimuli (Fig. 2d).

GCaMP6 performance was compared to other sensors in several ways. The fraction of responding neurons detected with GCaMP6s was threefold higher than for GCaMP5G (fivefold higher than GCaMP3) (Fig. 2e). Notably, the fractions of active neurons detected with GCaMP6s and GCaMP6m were also significantly higher than for OGB1-AM (Fig. 2e, f, $P < 0.01$, Wilcoxon rank sum test). GCaMP6 sensors thus reveal neuronal dynamics that were previously undetectable with protein sensors.

We imaged GCaMP6s-expressing neurons through a chronic imaging window^{6,31} over several weeks in V1 (Supplementary Fig. 6c)¹⁸. Sensory responses and orientation tuning of neurons were largely stable (Supplementary Fig. 6c, e). Responses of a small fraction of neurons were lost over time, which was balanced by other neurons that started responding. The orientation preference of the neurons responding in all imaging sessions (~56% of cells that responded in the first session) was remarkably stable (Supplementary Fig. 6d). Over months of expression a small fraction of highly expressing neurons acquired nuclear fluorescence; these neurons eventually also developed aberrant responses¹¹ (Supplementary Fig. 7). These experiments indicate that expression of GCaMP6s over 1–2 months does not obviously perturb the function of cortical circuits.

We directly compared cellular fluorescence changes and spiking using loose-seal, cell-attached recordings. The contrast of the visual stimulus was adjusted online to maintain a moderate spike rate. GCaMP6s

produced large fluorescence transients even in response to single action potentials (>6 times larger than for GCaMP5K, Fig. 3 and Supplementary Video 1), yielding high detection rates for single spikes ($99 \pm 0.2\%$; at 1% false-positive rate, $n = 9$ cells, 250 spikes). GCaMP6f and GCaMP6m showed slightly lower spike detection efficiencies, but with faster kinetics (Fig. 3, Supplementary Table 3). Individual spikes within a burst resulted in stepwise fluorescence increases (Fig. 3b), which were resolvable if they were separated by an interval on the order of the rise time of the sensor or more (100–150 ms, GCaMP6s; 75–100 ms, GCaMP6m; 50–75 ms, GCaMP6f; Supplementary Fig. 10, Supplementary Table 3). These data show that GCaMP6s can detect activity with near 100% action potential detection in pyramidal cells.

Calcium transients in dendritic spines

Pyramidal neuron output is shaped by thousands of excitatory synapses distributed across the dendritic arbor. Activation of single excitatory synapses causes calcium accumulations in individual dendritic spines, mediated by NMDARs^{32,33}, which can be imaged to measure the tuning of single synapses *in vivo*^{10,33}. We used GCaMP6s to image synaptic calcium signals within dendritic spines over chronic timescales. In V1 with sparsely labelled L2/3 pyramidal neurons (Methods) small dendritic branches were imaged at high magnification during visual stimulation (Fig. 4a). We first focused on neurons that did not fire visually evoked action potentials (~40% of neurons) to avoid calcium changes caused by action potentials back-propagating into dendrites⁷. Individual spines showed large fluorescence transients, often independent of their parent dendrites (Fig. 4b, d and Supplementary Video 2). Spine responses were orientation-tuned (Fig. 4b–e), as expected from the fact that most input to V1 neurons originates from (presumably orientation-tuned) V1 neurons³⁴. Neighbouring spines were often tuned with different preferred orientations. The correlation between orientation tuning and distance

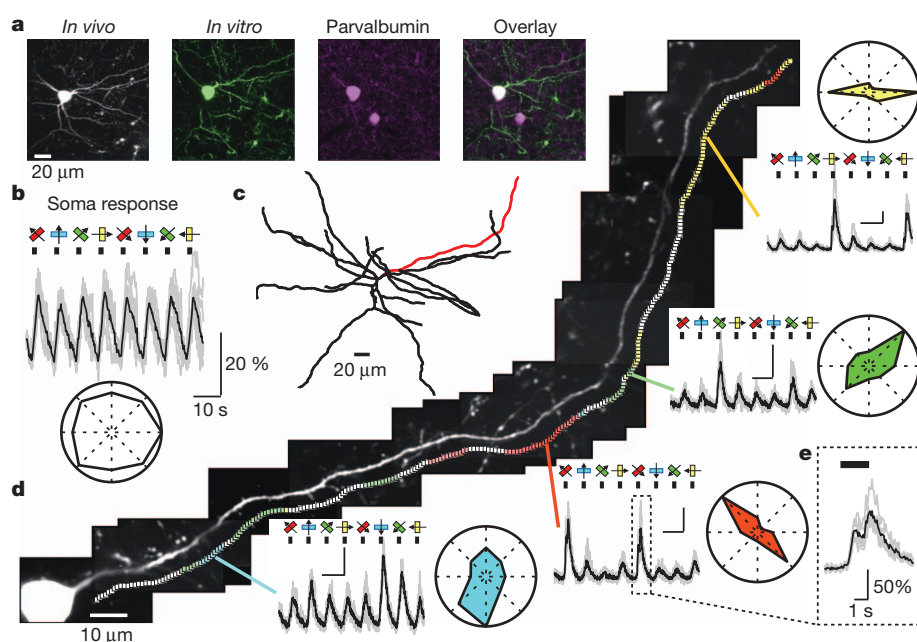


Figure 6 | Orientation-tuned domains in dendrites of GABAergic interneurons. **a**, A GCaMP6s-expressing interneuron (soma depth, 250 μ m), identified post hoc as a parvalbumin-positive interneuron. **b**, Somatic fluorescence changes to oriented drifting grating (same cell as in **a**). Bottom, polar plot. **c**, Reconstruction of the dendritic arbor based on GCaMP6s fluorescence. **d**, Left, a dendrite of the cell (red in **c**) was imaged along its entire length. Coloured squares indicate dendritic sites showing significant orientation tuning ($P < 0.01$, ANOVA across 8 stimulus directions). The colour of each square indicates the local preferred orientation, and the saturation of the colour encodes the orientation selectivity index (OSI = 1, saturated colour; OSI = 0, white). Right, example dendritic fluorescence changes and the corresponding polar plots for four locations with distinct orientation preferences. Scale bars, 10 s; 50% $\Delta F/F_0$. **e**, Zoomed-in view of the dendritic calcium signal corresponding to the box in **d**. The signal shows modulation at the frequency of the drifting grating (1 Hz).

between spines was weak ($R = 0.08$; $P > 0.05$). Overall, 27% of spines (62/228, 15 dendrites, 4 mice) were visually responsive, and most were orientation-tuned (Fig. 4f, g). Furthermore, close to 40% of spines were active at some time during the imaging session (Fig. 4f). Bleaching of GCaMP6 was negligible over 40 imaging trials (320 s of continuous imaging, Fig. 4h).

We imaged the same dendritic segments over imaging sessions separated by weeks. Most spines persisted over this timescale, although other spines appeared and disappeared³¹. We analysed the visual responses of persistent spines. The percentage of spines that responded to visual stimulation was stable over a week (37/139 on day 1; 34/139 on day 8). Spines that responded during one imaging session were also likely to respond a week later (65%). Furthermore, visually responsive spines mostly retained their orientation tuning (Fig. 4i–k).

How orientation-tuned neurons connect to other orientation-tuned neurons remains controversial^{10,35}. Some measurements suggest that individual neurons sample heterogeneous input, from neurons with diverse orientation-tuning¹⁰. Models indicate that specific connectivity is not required for orientation tuning³⁶. However, other measurements emphasize that neurons prefer to make synapses with neurons sharing similar orientation tuning³⁵, but orientation tuning at the level of sub-threshold membrane potential is weak^{37,38}. We thus compared the orientation tuning of populations of individual spines and the output of their parent neuron. We identified neurons with tuned output (OSI, 0.91 ± 0.04 , $n = 5$) (Fig. 5a) and measured the orientation tuning of large numbers of dendritic spines per neuron (average, 201 spines; range, 120–298 spines) (Fig. 5b–d). Contributions to the signal from back-propagating action potentials were removed using a computational subtraction procedure (Methods; Supplementary Fig. 11). For individual neurons, the orientation tuning averaged across all spines was biased towards the orientation tuning of the parent neurons (Fig. 5e, f) ($P < 0.01$, Wilcoxon rank sum test), although the modulation depth for spines was smaller ($P < 0.01$, Wilcoxon rank sum test). A similar trend was also apparent in the distributions of preferred orientations across spines (Fig. 5g, h). Our results show that spine fluorescence transients, which are not necessarily good predictors of the strength of excitatory synaptic input³⁹, averaged across a neuron predict the orientation tuning of the cell's output.

Calcium transients in GABAergic dendrites

Consistent with previous results^{40,41}, somatic fluorescence changes in GABAergic neurons were broadly tuned with respect to stimulus

orientation (Fig. 6a, b; Supplementary Fig. 12). However, their dendrites showed pronounced orientation-tuned domains (Fig. 6c, d). Individual dendritic branches often had multiple domains with distinct preferred orientations, consistent with the diverse preferences of inputs to GABAergic neurons^{42,43}. The underlying fluorescence responses were modulated at the temporal frequency of the drifting grating (1 Hz) (Fig. 6e), a characteristic feature of V1 excitatory neurons in the superficial layers of V1 (ref. 30), suggesting that the responses reflect local excitatory synaptic input. Tuned dendritic domains were seen in parvalbumin-positive (Fig. 6) and somatostatin-positive (data not shown) interneurons.

We mapped the visual response along 104 interneuron dendrites (total imaged length 6.3 mm) with regions of interest placed every micrometre along the dendrite (Fig. 6d). Visually-evoked GCaMP6s responses were detected in 5.7 mm (90%) of dendritic length, with 1.79 mm (28%) showing significant orientation-tuning ($P < 0.01$, ANOVA across 8 conditions). The majority of imaged dendrites (33/54; longer than 50 μ m) had multiple domains preferring different orientations. The OSI (0.44 ± 0.15 ; mean \pm s.d.) of tuned dendritic segments was higher than for interneuron somata ($P < 0.005$, Supplementary Fig. 12), but lower than for typical pyramidal neurons ($P < 0.001$, Supplementary Fig. 9). Back-propagation⁴⁴ of broadly tuned somatic action potentials probably lowers the OSI of the dendritic calcium signals. Consistent with this, both the dendritic OSI and the percentage of orientation selective sites increased with distance from the soma (Supplementary Fig. 13a).

The sizes of individual domains (12 ± 7 μ m, mean \pm s.d. full-width at half maximum, $n = 107$ domains) were considerably larger than the spacing between excitatory synapses on interneuron dendrites (< 1 μ m)⁴⁵ (Supplementary Fig. 13b, c). Nearby domains with distinct preferred orientations often overlapped (Supplementary Fig. 13b). The large size of these domains might reflect spatially clustered input with shared orientation preference. Alternatively, individual domains might be dominated by a few sparse, randomly distributed strong inputs⁴⁶, amplified by local postsynaptic mechanisms⁴⁷.

Conclusions

We developed a new class of genetically encoded indicators of neuronal function with greatly improved properties. GCaMP6 indicators cross important performance thresholds. They have higher sensitivity than commonly used synthetic calcium dyes (for example, OGB1-AM) and detect individual action potentials with high reliability at

reasonable microscope magnifications. These indicators can be used to image large groups of neurons as well as tiny synaptic compartments over multiple imaging sessions separated by months. It is likely that these sensors will find widespread applications for diverse problems in brain research and calcium signalling. Future engineering efforts could focus on red fluorescent calcium indicator proteins^{48,49}, which promise imaging considerably deeper in scattering tissue⁵⁰.

METHODS SUMMARY

GCaMP variants were tested in dissociated rat hippocampal neurons transduced using lentivirus-mediated gene transfer. *In vivo* imaging was performed in adult, anaesthetized C57BL/6 mice. Visual cortex neurons were transduced with GCaMP variants using AAV. Imaging experiments started 2–3 weeks after AAV injection.

Full Methods and any associated references are available in the online version of the paper.

Received 10 December 2012; accepted 4 June 2013.

- Baker, P. F., Hodgkin, A. L. & Ridgway, E. B. Depolarization and calcium entry in squid giant axons. *J. Physiol. (Lond.)* **218**, 709–755 (1971).
- Kerr, R. *et al.* Optical imaging of calcium transients in neurons and pharyngeal muscle of *C. elegans*. *Neuron* **26**, 583–594 (2000).
- Tank, D. W., Sugimori, M., Connor, J. A. & Llinas, R. R. Spatially resolved calcium dynamics of mammalian Purkinje cells in cerebellar slice. *Science* **242**, 773–777 (1988).
- Sabatini, B. L., Oertner, T. G. & Svoboda, K. The life cycle of Ca^{2+} ions in dendritic spines. *Neuron* **33**, 439–452 (2002).
- Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P. & Reid, R. C. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature* **433**, 597–603 (2005).
- Huber, D. *et al.* Multiple dynamic representations in the motor cortex during sensorimotor learning. *Nature* **484**, 473–478 (2012).
- Svoboda, K., Denk, W., Kleinfeld, D. & Tank, D. W. *In vivo* dendritic calcium dynamics in neocortical pyramidal neurons. *Nature* **385**, 161–165 (1997).
- Petreaanu, L. *et al.* Activity in motor-sensory projections reveals distributed coding in somatosensation. *Nature* **489**, 299–303 (2012).
- Wang, Y. *et al.* Stereotyped odor-evoked activity in the mushroom body of *Drosophila* revealed by green fluorescent protein-based Ca^{2+} imaging. *J. Neurosci.* **24**, 6507–6514 (2004).
- Jia, H., Rochefort, N. L., Chen, X. & Konnerth, A. Dendritic organization of sensory input to cortical neurons *in vivo*. *Nature* **464**, 1307–1312 (2010).
- Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).
- Zariwala, H. A. *et al.* A Cre-dependent GCaMP3 reporter mouse for neuronal imaging *in vivo*. *J. Neurosci.* **32**, 3131–3141 (2012).
- Mao, T., O'Connor, D. H., Scheuss, V., Nakai, J. & Svoboda, K. Characterization and subcellular targeting of GCaMP-type genetically-encoded calcium indicators. *PLoS One* **3**, e1796 (2008).
- Dreosti, E., Odermatt, B., Dorostkar, M. M. & Lagnado, L. A genetically encoded reporter of synaptic activity *in vivo*. *Nature Methods* **6**, 883–889 (2009).
- Guerrero, G. *et al.* Heterogeneity in synaptic transmission along a *Drosophila* larval motor axon. *Nature Neurosci.* **8**, 1188–1196 (2005).
- Akerboom, J. *et al.* Optimization of a GCaMP calcium indicator for neural activity imaging. *J. Neurosci.* **32**, 13819–13840 (2012).
- Nakai, J., Ohkura, M. & Imoto, K. A high signal-to-noise Ca^{2+} probe composed of a single green fluorescent protein. *Nature Biotechnol.* **19**, 137–141 (2001).
- Mank, M. *et al.* A genetically encoded calcium indicator for chronic *in vivo* two-photon imaging. *Nature Methods* (2008). **5**, 805–811.
- Miyawaki, A. *et al.* Fluorescent indicators for Ca^{2+} based on green fluorescent proteins and calmodulin. *Nature* **388**, 882–887 (1997).
- Nagai, T., Yamada, S., Tominaga, T., Ichikawa, M. & Miyawaki, A. Expanded dynamic range of fluorescent indicators for Ca^{2+} by circularly permuted yellow fluorescent proteins. *Proc. Natl Acad. Sci. USA* **101**, 10554–10559 (2004).
- Palmer, A. E. *et al.* Ca^{2+} indicators based on computationally redesigned calmodulin-peptide pairs. *Chem. Biol.* **13**, 521–530 (2006).
- Horikawa, K. *et al.* Spontaneous network activity visualized by ultrasensitive Ca^{2+} indicators, yellow Cameleon-Nano. *Nature Methods* **7**, 729–732 (2010).
- Pologruto, T. A., Yasuda, R. & Svoboda, K. Monitoring neural activity and $[\text{Ca}^{2+}]$ with genetically encoded Ca^{2+} indicators. *J. Neurosci.* **24**, 9572–9579 (2004).
- Reiff, D. F. *et al.* *In vivo* performance of genetically encoded indicators of neural activity in flies. *J. Neurosci.* **25**, 4766–4778 (2005).
- Baird, G. S., Zacharias, D. A. & Tsien, R. Y. Circular permutation and receptor insertion within green fluorescent proteins. *Proc. Natl Acad. Sci. USA* **96**, 11241–11246 (1999).
- Crivici, A. & Ikura, M. Molecular and structural basis of target recognition by calmodulin. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 85–116 (1995).
- Akerboom, J. *et al.* Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *J. Biol. Chem.* **284**, 6455–6464 (2009).
- Bayley, P. M., Findlay, W. A. & Martin, S. R. Target recognition by calmodulin: dissecting the kinetics and affinity of interaction using short peptide sequences. *Protein Sci.* **5**, 1215–1228 (1996).
- Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnol.* **22**, 1567–1572 (2004).
- Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).
- Trachtenberg, J. T. *et al.* Long-term *in vivo* imaging of experience-dependent synaptic plasticity in adult cortex. *Nature* **420**, 788–794 (2002).
- Mainen, Z. F., Malinow, R. & Svoboda, K. Synaptic calcium transients in single spines indicate that NMDA receptors are not saturated. *Nature* **399**, 151–155 (1999).
- Chen, X., Leischner, U., Rochefort, N. L., Nelken, I. & Konnerth, A. Functional mapping of single spines in cortical neurons *in vivo*. *Nature* **475**, 501–505 (2011).
- Binzegger, T., Douglas, R. J. & Martin, K. A. A quantitative map of the circuit of cat primary visual cortex. *J. Neurosci.* **24**, 8441–8453 (2004).
- Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
- Hansel, D. & van Vreeswijk, C. The mechanism of orientation selectivity in primary visual cortex without a functional map. *J. Neurosci.* **32**, 4049–4064 (2012).
- Liu, B. H. *et al.* Broad inhibition sharpens orientation selectivity by expanding input dynamic range in mouse simple cells. *Neuron* **71**, 542–554 (2011).
- Tan, A. Y., Brown, B. D., Scholl, B., Mohanty, D. & Priebe, N. J. Orientation selectivity of synaptic input to neurons in mouse and cat primary visual cortex. *J. Neurosci.* **31**, 12339–12350 (2011).
- Sobczyk, A., Scheuss, V. & Svoboda, K. NMDA receptor subunit-dependent $[\text{Ca}^{2+}]$ signaling in individual hippocampal dendritic spines. *J. Neurosci.* **25**, 6037–6046 (2005).
- Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K. & Tsumoto, T. GABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by *in vivo* functional Ca^{2+} imaging in transgenic mice. *J. Neurosci.* **27**, 2145–2149 (2007).
- Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* **67**, 858–871 (2010).
- Bock, D. D. *et al.* Network anatomy and *in vivo* physiology of visual cortical neurons. *Nature* **471**, 177–182 (2011).
- Hofer, S. B. *et al.* Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neurosci.* **14**, 1045–1052 (2011).
- Goldberg, J. H., Tamas, G. & Yuste, R. Ca^{2+} imaging of mouse neocortical interneurone dendrites: la-type K^{+} channels control action potential backpropagation. *J. Physiol. (Lond.)* **551**, 49–65 (2003).
- Gulyás, A. I., Megias, M., Emri, Z. & Freund, T. F. Total number and ratio of excitatory and inhibitory synapses converging onto single interneurons of different types in the CA1 area of the rat hippocampus. *J. Neurosci.* **19**, 10082–10097 (1999).
- Goldberg, J. H., Tamas, G., Aronov, D. & Yuste, R. Calcium microdomains in aspiny dendrites. *Neuron* **40**, 807–821 (2003).
- Katona, G. *et al.* Roller Coaster Scanning reveals spontaneous triggering of dendritic spikes in CA1 interneurons. *Proc. Natl Acad. Sci. USA* **108**, 2148–2153 (2011).
- Zhao, Y. *et al.* An expanded palette of genetically encoded Ca^{2+} indicators. *Science* **333**, 1888–1891 (2011).
- Akerboom, J. *et al.* Genetically encoded calcium indicators for multi-color neural activity imaging and combination with optogenetics. *Front. Mol. Neurosci.* **6**, 2 (2013).
- Kobat, D. *et al.* Deep tissue multiphoton microscopy using longer wavelength excitation. *Opt. Express* **17**, 13354–13364 (2009).
- Wang, Q., Shui, B., Kotlikoff, M. I. & Söndermann, H. Structural basis for calcium sensing by GCaMP2. *Structure* **16**, 1817–1827 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Akerboom and L. Tian for constructs and advice; J. Hasseman, M. Ramirez, G. Tsegaye for molecular cloning; B. Shields for neuronal culture; A. Hu for histology; B. Fosque, R. Behnam, K. Ritola for virus production; J. Macklin and R. Patel for spectroscopy; B. Coop and L. Ramasamy for multiwell electrode manufacturing; K. Smith for mouse viral transduction; K. Hibbard for fly husbandry; and J. Yu, C. Niell, M. Stryker, J. Trachtenberg and A. Kerlin for advice on visual cortex experiments. S.L.R. is supported by a Fellowship from the Swiss National Science Foundation. M.B.O. is supported by a Marie Curie Career Integration Grant PCIG09-GA-2011-294049.

Author Contributions V.J., R.A.K., L.L.L. and K.S. initiated the project. T.-W.C., T.J.W. and D.S.K. conducted neuronal culture screening. T.-W.C. performed mouse visual cortical experiments on pyramidal neurons. T.-W.C. and A.B. performed experiments on inhibitory neurons. Y.S., T.J.W. and S.R.P. carried out fly larval neuromuscular junction studies. Y.S. and V.J. carried out adult fly antennal lobe imaging. S.L.R. and M.B.O. conducted zebrafish tectal imaging. E.R.S. performed protein assays. All authors analysed data. T.-W.C., R.A.K., M.B.O., V.J., L.L.L., K.S. and D.S.K. wrote the paper with comments from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.B.O. (michael.orger@neuro.fchamplimaud.org) for zebrafish, V.J. (jayaramanv@janelia.hhmi.org) for flies, L.L.L. (loogerl@janelia.hhmi.org) for GCaMP protein structure information, K.S. (svoboda@janelia.hhmi.org) for mice and D.S.K. (kimd@janelia.hhmi.org) for neuronal culture screen information and constructs.

METHODS

All experiments were conducted according to National Institutes of Health guidelines for animal research and were approved by the Janelia Farm Research Campus Institutional Animal Care and Use Committee and Institutional Biosafety Committee.

Neuronal culture screen. GCaMP variants were made in a modified SIV-based lentiviral construct, pGP-*syn*-GCaMP-nls-mCherry-WPRE, derived from pCL20cSLFR MSCV-GFP⁵². The proviral vector included a 476 base pair human synapsin promoter, GCaMP, a nuclear localization sequence fused to mCherry, and the woodchuck hepatitis post-transcriptional regulatory element. Site-directed mutagenesis was conducted by PCR and mutated regions were incorporated into the lentiviral constructs by gene assembly⁵³.

Hippocampi were dissected and dissociated in papain. Cells were plated at a density of 225,000 viable cells per well in 24-well glass-bottom plates (Mattek, #1.5 glass coverslips), pre-coated with Matrigel (BD Biosciences). Cells were cultured in growth medium (28 mM glucose, 2.4 mM sodium bicarbonate, 100 $\mu\text{g ml}^{-1}$ transferrin, B-27 supplement (1 \times , Invitrogen), 500 μM L-glutamine, 50 units ml^{-1} penicillin, 50 $\mu\text{g ml}^{-1}$ streptomycin, 5% fetal bovine serum in MEM).

Lentiviral particles were made in a biosafety level 2 laboratory by transfecting a proviral construct and packaging and coat pseudotyping DNA constructs (pCAG-SIVgprre, pCAG4-RTR-SIV, pCMV-VSV-G)^{52,54} into HEK293T/17 cells (ATCC) in 10-cm plates. After 72 h, supernatant was collected (6 ml) and filtered. Neuronal cultures were infected at 3 days *in vitro*. Each well of a 24-well plate was incubated overnight with 0.5 ml of lentivirus in conditioned growth medium. The growth medium was supplemented with 4 μM AraC to inhibit glial proliferation. In some experiments, OGB1-AM was loaded into cells by incubating neurons in 1 ml of 2 μM OGB1-AM (Invitrogen) for 30 min and rinsing 3 times with imaging buffer (145 mM NaCl, 2.5 mM KCl, 10 mM glucose, 10 mM HEPES, pH 7.4, 2 mM CaCl_2 , 1 mM MgCl_2).

Neurons were stimulated in imaging buffer containing a drug cocktail to inhibit synaptic receptors (10 μM CNQX, 10 μM (R)-CPP, 10 μM gabazine, 1 mM (S)-MCPG, Tocris Bioscience). Under these conditions, intracellular calcium increases are presumably caused by the opening of voltage sensitive calcium channels.

Action potentials (APs) (83 Hz) were evoked by field stimulation with a Grass Technologies S48 stimulation unit and a custom-built 24-well cap stimulator with pairs of parallel platinum wires. The microscope was an Olympus IX81 with a $\times 10$ (0.4 NA) air objective lens and EMCCD camera (Andor 897, 512 \times 512 pixels, 35 frames per second), Cairn OptoLED illumination system, and GFP (Excitation: 450–490 nm; Dichroic: 495 nm long-pass; Emission: 500–550 nm) and TxRed (Excitation: 540–580 nm; Dichroic: 585 nm long-pass; Emission: 593–668 nm) filter sets. The field of view was 800 $\mu\text{m} \times$ 800 μm . Images were background subtracted (mean of 5% lowest pixel values). Responses were quantified for each cell as change in fluorescence divided by baseline fluorescence measured one second before stimulation. Signal-to-noise ratio (SNR) was quantified as peak $\Delta F/F_0$ response over the standard deviation of the signal during a one second period before stimulation.

Control experiments varying stimulation voltage, frequency and pulse width insured suprathreshold stimulation of neurons. Voltage imaging using the ArchWT-GFP archaerhodopsin-based voltage sensor⁵⁵ confirmed that individual pulses (1 ms, 40 V, 83 Hz) reliably triggered single APs. The imaging and stimulation system was controlled by custom scripts written in MetaMorph software (version 7.7.5, Molecular Devices) and Ephus software⁵⁶ (<http://www.ephus.org>). Detailed neuronal culture screening methods will be described elsewhere (T.J.W., T.W.C., E.R.S., R.A.K., V.J., L.L.L., K.S. and D.S.K., manuscript in preparation).

Labelling V1 neurons. Constructs used to produce AAV included pGP-AAV-*syn*-GCaMP-WPRE and the Cre recombinase-activated construct pGP-AAV-*syn*-flex-GCaMP-WPRE. Virus was injected slowly (30 nl in 5 min) at a depth of 250 μm into the primary visual cortex (two sites, 2.5 and 2.9 mm lateral from the lambda suture). For population imaging and electrophysiology (Figs 2 and 3), AAV2/1-*syn*-GCaMP-WPRE virus (titer: $\sim 10^{11}$ to 10^{12} genomes ml^{-1}) was injected into the visual cortex of C57BL/6J mice (1.5–2 months old)⁶. For dendritic imaging (Figs 4–6), sparse labelling was achieved by injecting a mixture of diluted AAV2/1-*syn*-Cre particles (titer: $\sim 10^{12}$ genomes ml^{-1} , diluted 8,000–20,000-fold in PBS) and high titre, Cre-dependent GCaMP6s virus ($\sim 8 \times 10^{11}$ genomes ml^{-1}). This produces strong GCaMP6 expression in a small subset of neurons (~ 3 –5 cells in a 250 $\mu\text{m} \times$ 250 μm volume), defined by Cre expression⁵⁷. Both pyramidal (Figs 4 and 5) and GABAergic (Fig. 6) neurons were labelled using this approach, but they could be distinguished based on the presence or absence of dendritic spines. Post hoc immunolabelling further identified the imaged cells. For specific labelling of parvalbumin interneurons (Supplementary Figs 12, 13), Cre-dependent GCaMP6s AAV was injected into the visual

cortex of PV-IRES-Cre mice⁵⁸. Individual somata (Supplementary Fig. 12) and dendritic segments could be recognized (Supplementary Fig. 13; total length of imaged dendrite: 2.86 mm), but the high labelling density made it difficult to track individual dendrites over long distances.

Window surgery. After 2–4 weeks of expression, mice were anaesthetized using isoflurane (3% for induction, 1.5–2% during surgery) and a circular craniotomy (2–3 mm diameter) was made above V1 (centred 2.7 mm lateral from the lambda suture). For acute experiments, the craniotomy was covered with agarose (1–1.3%), and a round glass coverslip (Warner Instruments; 5 mm diameter; #1 thickness) was cemented to the skull to reduce motion of the exposed brain. A custom titanium head post was fixed to the skull using black dental cement (Contemporary Ortho-Jet). For simultaneous imaging and cell-attached recording, the exposed brain was covered with ~ 1 mm thick agarose (1.3%) without a coverslip. For chronic imaging experiments, the imaging window was constructed from two layers of microscope coverglass⁶. A larger piece (Fisher, #1 thickness) was attached to the bone and a smaller insert (#2 thickness) was fitted snugly into the craniotomy. Imaging experiments were started ~ 1 –2 weeks after chronic window implantation.

Visual stimuli. Moving grating stimuli were generated using the Psychophysics Toolbox^{59,60} in MATLAB. Each stimulus trial consisted of a 4 s blank period (uniform grey at mean luminance) followed by a 4 s drifting sinusoidal grating (0.05 cycles per degree, 1 Hz temporal frequency). Typically, 8 drifting directions were used (separated by 45 degrees) and 5 trials were recorded for each direction, giving a total of 40 stimulus trials per recording session (320 s recording time). The gratings were presented with an LCD monitor (30 \times 40 cm), placed 25 cm in front of the centre of the right eye of the mouse. The monitor subtended an angle of $\pm 38^\circ$ horizontally and -20° to $+38^\circ$ vertically around the eye of the mouse. For experiments with cell-attached recording (Fig. 3), pipette access required the use of a smaller LCD monitor (12 \times 16 cm) placed ~ 10 cm in front of the right eye. During simultaneous imaging and electrophysiology, the optimal grating stimulus was repeatedly played (duration 2 s, separated by a 4 s blank period), but the contrast of the stimulus grating was adjusted online to maintain moderate spike rates.

For analysis of dendritic spine populations (Fig. 5), the stimulus contrast was lowered to reduce action potential-related dendritic signals (average $\Delta F/F_0$ in dendritic shafts at the preferred orientations was 32%, 80%, 16%, 15%, 12% for cells 1 through 5, corresponding to stimulus contrast 10–40%, 5–20%, 10–40%, 20–40%, 20–40%). The orientation preference of dendritic shafts was identical to the soma, consistent with back-propagating action potentials (data not shown). Orientation tuning of GABAergic dendrites (Fig. 6) was mapped using the standard stimulus set at full contrast. The stimulus duration was 2 s; because of the slower decay of calcium transients⁴¹, we used a 6 s inter-trial interval.

Imaging mouse V1 neurons and dendrites. Mice were placed on a warm blanket (37 $^\circ\text{C}$) and kept anaesthetized with 0.5% isoflurane and sedated with chlorprothixene (20–40 μl at 0.33 mg ml^{-1} , intramuscular (i.m.))³⁰. Imaging was performed using a custom-built two-photon microscope (designs available at (<http://research.janelia.org/Svoboda>) equipped with a resonant galvo scanning module (Thorlabs), controlled by ScanImage (<http://scanimage.org>)⁶¹. The light source was a Mai Tai femtosecond pulsed laser (Spectra-Physics) running at 940 nm. The objective was a $\times 16$ water immersion lens (Nikon, 0.8 NA, 3 mm working distance). The power used was 35–50 mW for full field imaging (Fig. 2) and 20–40 mW for higher zoom imaging (Figs 3–6).

Images were collected at 15 Hz (512 \times 512 pixels, 250 $\mu\text{m} \times$ 250 μm ; Fig. 2) or 60 Hz (256 \times 256 pixels, 30 $\mu\text{m} \times$ 30 μm ; Fig. 3), or 15 Hz (512 \times 512 pixels, 30 $\mu\text{m} \times$ 30 μm ; Figs 4 and 5), or 15 Hz (512 \times 512 pixels, 30 $\mu\text{m} \times$ 30 μm to 100 $\mu\text{m} \times$ 100 μm ; Fig. 6). For dendritic imaging experiments (Figs 4–6), fields of view were chosen so that extended dendritic segments were in one focal plane. At the end of each imaging session, z-stacks (1 μm step size) of the recorded cells were acquired. The coordinates of the imaged dendrites relative to the parent somata were recorded. The orientation, curvature and the branching pattern of the dendrites together with the constellation of spines, helped to precisely identify the same field of view in long-term imaging experiments.

Electrophysiology. *In vivo* cell-attached recordings were performed using glass pipettes (~ 5 –7 M Ω) filled with solution containing the following (in mM): 125 NaCl, 5 KCl, 10 glucose, 10 HEPES, 2 CaCl_2 , 2 MgSO_4 , and 0.1 Alexa Fluor 594; pH 7.4). Signals were amplified using an AxoPatch 200B amplifier (Molecular Devices), filtered at 5 kHz, and digitized at 10 kHz. Spikes were recorded using current clamp mode. The frame trigger pulses of ScanImage 4.0 were also recorded and used offline to synchronize individual frames to electrophysiological recordings. After establishment of a low-resistance seal (15–50 M Ω), the orientation, spatial and temporal frequency of the stimuli was quickly optimized for individual neurons using recorded spikes. The optimal grating stimulus was repeated at a reduced contrast to maintain a moderate spiking rate.

Image analysis. Mechanical drift in the imaging plane was corrected using the TurboReg plug-in in ImageJ⁶². All remaining analyses were performed in MATLAB. Regions of interest (ROIs) corresponding to visually identifiable cell bodies were selected using a semi-automated algorithm (Supplementary Fig. 14). For GCaMP, ring-shaped ROIs were placed at the cytosolic regions of the cells (excluding the nucleus; GCaMP expression is typically restricted to the cytoplasm¹¹). For OGB1-AM, circular ROIs covering the whole soma were used. For long-term GCaMP imaging, baseline fluorescence images of multiple sessions were inspected manually, and only the cells that could be clearly identified in all imaged sessions were included in the analysis. The fluorescence time course of each cell was measured by averaging all pixels within the ROI, with a correction for neuropil contamination⁴¹. The fluorescence signal of a cell body was estimated as $F_{\text{cell_true}}(t) = F_{\text{cell_measured}}(t) - r \times F_{\text{neuropil}}(t)$, with $r = 0.7$. The neuropil signal $F_{\text{neuropil}}(t)$ surrounding each cell was measured by averaging the signal of all pixels within a 20- μm region from the cell centre (excluding all selected cells). Cell-attached recordings confirmed that neuropil-compensated fluorescence changes reflect action potentials in single neurons (Supplementary Fig. 15). To ensure robust neuropil subtraction, only cells that were at least 3% brighter than the surrounding neuropil were included. The neuropil correction was not applied for dendritic imaging experiments because sparse labelling provided negligibly low background. $\Delta F/F_0$ was calculated as $(F - F_0)/F_0$, where F_0 is the baseline fluorescence signal averaged over a 2-s period immediately before the start of visual stimulation. Visual responses were measured for each trial as $\Delta F/F_0$, averaged over the stimulus period. Visually responsive neurons were defined as cells with significant stimulus-related fluorescence changes (ANOVA across blank and eight direction periods, $P < 0.01$)⁵ with an average $\Delta F/F_0$ at preferred orientations greater than 6%.

The orientation selectivity index (OSI) was calculated for visually responsive cells^{16,30}. First, the preferred orientation (θ_{pref}) of the cell was determined as the stimulus that produced the strongest response. The orientation tuning curve was constructed by measuring the mean $\Delta F/F_0$, averaged over the stimulus period, for each orientation. We then fitted the tuning curve with the sum of two Gaussians centred on θ_{pref} and $\theta_{\text{pref}} + \pi$, both with width σ (constrained to $>15^\circ$), amplitudes A_1 and A_2 , and a constant baseline B . The OSI was defined as $\text{OSI} = (R_{\text{pref}} - R_{\text{ortho}})/(R_{\text{pref}} + R_{\text{ortho}})$, where R_{pref} and R_{ortho} are the response amplitudes at the preferred (θ_{pref}) and the orthogonal orientation ($\theta_{\text{pref}} + \pi/2$) respectively.

For simultaneous imaging and cell-attached recording, ring-shaped ROIs were placed over the cytosolic regions of the cells. Fluorescence transients at the soma were caused by action potentials, with little contribution from subthreshold activity⁶³ (Supplementary Fig. 15). To quantify the efficiency for detecting single APs (Fig. 3), we identified single AP events with nearby APs at least 1 s away. Fluorescence traces consisting of 10 frames (0.17 s) before and 60 frames (1 s) after the i^{th} 1 AP event were assembled in 70-dimensional vectors, f_i . Segments of noisy traces, n_i , were taken from periods without APs. The average of all 1 AP traces was used as a template vector, $f_{\text{template}} = \sum_i f_i/N$. The vector was normalized after subtraction of the mean to create a unit vector $\hat{f}_{\text{template}}$. The projection of f_i or n_i along the direction of $\hat{f}_{\text{template}}$ was calculated to obtain a scalar f_i or n_i , respectively. The AP detection threshold was defined as the 99th percentile of all n_i values (that is, 1% false positive), and the percentage of the f_i values above the detection threshold was the AP detection efficiency.

For spine images (Figs 4 and 5), circular ROIs were placed over individual dendritic spines to measure spine fluorescence and compute $\Delta F/F_{0_spine}$. To minimize contamination from back-propagating action potentials (BAPs), we either recorded from 'silent cells' ($\sim 40\%$ of cells) showing few or no APs in response to a standard set of grating stimuli (Fig. 4), or used stimuli with reduced stimulus contrast (Fig. 5).

Occasional BAP related calcium signals that invaded the imaged spines were removed using a subtraction method, implemented in three steps (Supplementary Fig. 11). First, a region covering the entire parent dendritic shaft ($\sim 30 \mu\text{m}$ of dendritic length; excluding all spines) was drawn for each recorded dendritic segment to estimate BAP related global dendritic signal, $\Delta F/F_{0_dendrite}$. Because of the much larger volume of the dendritic shaft compared to tuned dendritic spines (100-fold), spines above and below the focal plane were expected to contribute negligible signal to $\Delta F/F_{0_dendrite}$. This was verified using principle component analysis (data not shown). Plotting $\Delta F/F_{0_spine}$ against $\Delta F/F_{0_dendrite}$ reveals two components of spine signals, a BAP-related component and a spine-specific component. Second, the BAP-related component was removed from the spine signals by subtracting a scaled version of the dendritic shaft signal, $\Delta F/F_{0_spine_specific} = \Delta F/F_{0_spine} - \alpha \Delta F/F_{0_dendrite}$. α was determined using robust regression (MATLAB function 'robustfit.m') of $\Delta F/F_{0_spine}$ vs. $\Delta F/F_{0_dendrite}$ (the slope of the fitted line in Supplementary Fig. 11b). Third, the visual responsiveness ($\Delta F/F_0 > 10\%$) and the OSI of individual spines were calculated with the BAP signal removed. Active spines (Fig. 4f) were defined as spines showing at least three spine-specific (that is, BAP independent) calcium events during the 5-min imaging session, with an event defined as an episode of the calcium signal that crosses 3 s.d. of the baseline noise for at least three consecutive frames (~ 50 ms).

We next confirmed the effectiveness of the BAP removal algorithm. First, BAP removed spine signals showed sharp orientation tuning ($\text{OSI} = 0.84 \pm 0.14$, mean \pm s.d., $n = 190$ spines) comparable to V1 excitatory neurons (Supplementary Fig. 9, $P > 0.05$, Wilcoxon rank sum). This holds even for spines preferring the orthogonal orientation compared to the soma ($\text{OSI} = 0.82 \pm 0.16$, $n = 24$ spines), indicating that the BAP contamination was cleanly removed. Second, the majority (79.7%) of visually responsive spines showed little trial-to-trial correlation with the dendritic shaft signal after BAP subtraction. 20.3% spines still displayed significant correlation with shaft responses compared to trial shuffled controls ($P < 0.01$). This could reflect synchronously active presynaptic cells, or imperfect BAP action potential signal subtraction. Because we were unable to distinguish between these possibilities, these spines were excluded from further analysis. In two cells (cell 4 and cell 5; Fig. 5), we collected larger numbers of trials (15 trials per orientation), which made it possible to further exclude trials with detectable dendritic responses ($\Delta F/F_0 > 6\%$). The result (that is, the preferred orientation of the summed spine responses) was identical as with using the subtraction procedure alone.

For the analysis of GABAergic cells (Fig. 6), dendrites were traced using 'Simple neurite tracer' in ImageJ⁶⁴. The program outputs a 1D sequence of coordinates traversing a dendrite and a 2D mask covering the traced dendrite, which were used to define ROIs along the dendrite (size, 1.5 μm of dendritic length; spacing, 1 μm). Visual responsiveness and orientation selectivity index were computed for individual ROIs. A region of dendrite is considered orientation selective if its response to at least one stimulus orientation is significantly different from other orientations ($P < 0.01$, ANOVA across 8 conditions). Because synaptic and BAP signals were intermixed in the same dendritic compartment, no attempt was made to isolate synaptic signals from BAP-related components. We did not analyse the relationship between integrated synaptic signals and the output tuning in GABAergic cells.

Reagent distribution. DNA constructs, AAV particles and *Drosophila* with GCaMP6 variants were deposited for distribution at Addgene (<http://www.addgene.org>), the University of Pennsylvania Vector Core (<http://www.med.upenn.edu/gtp/vectorcore>) and the Bloomington *Drosophila* Stock Center (<http://flystocks.bio.indiana.edu>), respectively.

52. Hanawa, H. *et al.* Efficient gene transfer into rhesus repopulating hematopoietic stem cells using a simian immunodeficiency virus-based lentiviral vector system. *Blood* **103**, 4062–4069 (2004).
53. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods* **6**, 343–345 (2009).
54. Stewart, S. A. *et al.* Lentivirus-delivered stable gene silencing by RNAi in primary cells. *RNA* **9**, 493–501 (2003).
55. Kralj, J. M., Douglass, A. D., Hochbaum, D. R., MacLaurin, D. & Cohen, A. E. Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. *Nature Methods* **9**, 90–95 (2012).
56. Suter, B. A. *et al.* Ephus: multipurpose data acquisition software for neuroscience experiments. *Front. Neural Circuits* **4**, 100 (2010).
57. Xu, N. L. *et al.* Nonlinear dendritic integration of sensory and motor input during an active sensing task. *Nature* **492**, 247–251 (2012).
58. Hippenmeyer, S. *et al.* A developmental switch in the response of DRG neurons to ETS transcription factor signaling. *PLoS Biol.* **3**, e159 (2005).
59. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
60. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* **10**, 437–442 (1997).
61. Polgarto, T. A., Sabatini, B. L. & Svoboda, K. ScanImage: Flexible software for operating laser-scanning microscopes. *Biomed. Eng. Online* **2**, 13 (2003).
62. Thévenaz, P., Ruttimann, U. E. & Unser, M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7**, 27–41 (1998).
63. Bandyopadhyay, S., Shamma, S. A. & Kanold, P. O. Dichotomy of functional organization in the mouse auditory cortex. *Nature Neurosci.* **13**, 361–368 (2010).
64. Longair, M. H., Baker, D. A. & Armstrong, J. D. Simple Neurite Tracer: open source software for reconstruction, visualization and analysis of neuronal processes. *Bioinformatics* **27**, 2453–2454 (2011).

Structure and function of the *Salmonella* Typhi chimaeric A₂B₅ typhoid toxin

Jeongmin Song^{1*}, Xiang Gao^{1*} & Jorge E. Galán¹

Salmonella enterica serovar Typhi (*S. Typhi*) differs from most other salmonellae in that it causes a life-threatening systemic infection known as typhoid fever¹. The molecular bases for its unique clinical presentation are unknown². Here we find that the systemic administration of typhoid toxin, a unique virulence factor of *S. Typhi*, reproduces many of the acute symptoms of typhoid fever in an animal model. We identify specific carbohydrate moieties on specific surface glycoproteins that serve as receptors for typhoid toxin, which explains its broad cell target specificity. We present the atomic structure of typhoid toxin, which shows an unprecedented A₂B₅ organization with two covalently linked A subunits non-covalently associated to a pentameric B subunit. The structure provides insight into the toxin's receptor-binding specificity and delivery mechanisms and reveals how the activities of two powerful toxins have been co-opted into a single, unique toxin that can induce many of the symptoms characteristic of typhoid fever. These findings may lead to the development of potentially life-saving therapeutics against typhoid fever.

Salmonella enterica serovar Typhi (*S. Typhi*), the cause of typhoid fever, results in more than 200,000 annual deaths^{1,3,4}. Unlike other *Salmonella* serovars, which typically cause self-limiting gastroenteritis, *S. Typhi* causes a systemic, life-threatening disease¹. The genome of *S. Typhi* contains very few unique virulence factors that are not found in non-typhoidal serovars, and the molecular bases for its unique virulence properties and clinical presentation are unknown⁵. One of the few *S. Typhi*-specific factors that have been shown to directly affect its interaction with host cells is an AB-type toxin dubbed typhoid toxin^{6–8}. Unlike typical AB toxins⁹, typhoid toxin is composed of two A subunits, PltA and CdtB, which are homologues of the A subunits of the pertussis and cytolethal distending toxins, respectively⁷. Its single B subunit, PltB, is a homologue of one of the components of the heteropentameric B subunit of pertussis toxin. Although the cellular targets of the ADP-ribosyl transferase activity of PltA have not yet been identified, CdtB is a DNase that inflicts DNA damage and induces cell-cycle arrest^{6,10,11}. *S. Typhi* produces typhoid toxin only within mammalian cells, and the toxin is then ferried to the extracellular environment by a unique transport mechanism that involves vesicle carrier intermediates^{7,8}. Convalescent typhoid fever patients developed strong immune response to the toxin components, indicating that typhoid toxin is produced during human infection^{12,13}. However, the potential contribution of typhoid toxin to disease and pathogenesis is unknown.

We found that systemic administration into mice of a highly purified preparation of biologically active typhoid toxin (Fig. 1a–c) caused many of the symptoms observed during the acute phase of typhoid fever^{1,14}. Despite the lack of fever (Supplementary Fig. 1), the mice appeared lethargic, showing clear signs of stupor and malaise (Supplementary Video 1), lost weight (Fig. 1d), and eventually died (Fig. 1e). The toxin-injected animals also showed a significant reduction in the number of circulating immune cells, resulting in the almost complete depletion of circulating neutrophils (Fig. 1f, g), a phenotype often observed during the acute phase of typhoid fever¹⁴. Consistent with this

observation, typhoid toxin was able to intoxicate a broad range of cells *in vitro*, including several epithelial and immune cells (Supplementary Fig. 2). Although symptoms were observed in animals inoculated with a toxin carrying a catalytic mutant of PltA (Fig. 1d), no detectable symptoms were observed when animals were inoculated with a toxin carrying a catalytic mutant of CdtB (Fig. 1d–g). Taken together, these results indicate that typhoid toxin, through its CdtB subunit, may contribute to the acute symptomatology observed during typhoid fever.

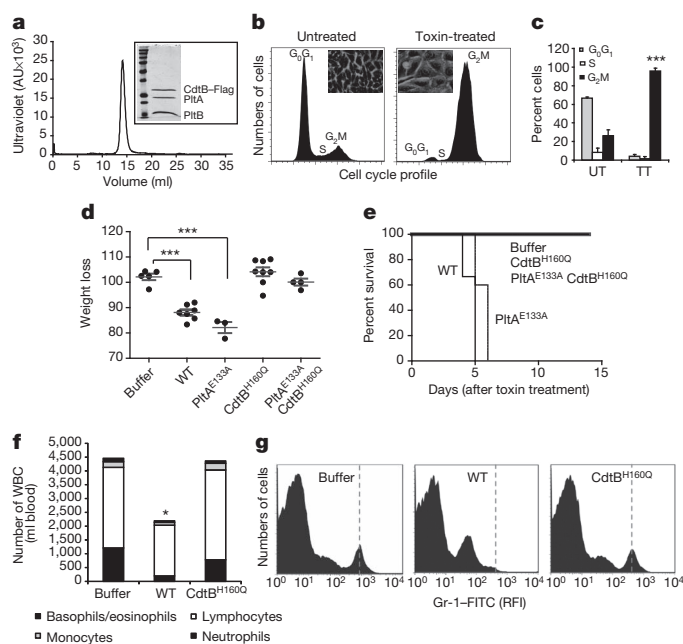


Figure 1 | Systemic administration of typhoid toxin causes symptoms observed during the acute phase of typhoid fever. **a**, Chromatographic profile of the typhoid toxin holotoxin used in the biological assays. The inset shows a Coomassie-blue-stained SDS-PAGE analysis of the peak fraction. AU, absorbance units. **b**, **c**, Typhoid toxin induces cell-cycle arrest in cultured cells. Human intestinal epithelial Henle-407 cells were left untreated, or treated with purified typhoid toxin and then analysed by flow cytometry. The insets show representative light microscope images of mock- or toxin-treated cells and is representative of at least three independent experiments. **c**, Averages of cell cycle profiles from at least three independent experiments. Bar represents average \pm standard deviation. *** $P < 0.001$, compared to the number of cells in G₂M of the control untreated group. UT, untreated; TT, toxin-treated. **d**, Weight loss 5 days after intravenous administration of different typhoid toxin preparations. Lines are the mean \pm standard error of the mean and represent the weight relative to the values before treatment. *** $P < 0.0001$. **e**, Survival of animals receiving different typhoid toxin preparations. $n = 3$ –5 animals per group. **f**, Circulating white blood cells were counted in a haematology analyser after the indicated treatments (* $P < 0.05$). **g**, Alternatively, the number of neutrophils (vertical dashed line) in peripheral blood of animals treated as indicated was determined by flow cytometry. RFI, relative fluorescence intensity.

¹Department of Microbial Pathogenesis, Yale University School of Medicine, New Haven, Connecticut 06536, USA.

*These authors contributed equally to this work.

To identify the cellular receptor(s) of typhoid toxin, we used a highly purified biologically active typhoid toxin preparation to affinity-purify biotin-labelled host cell-surface-interacting proteins. We found by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses of interacting proteins that, in human epithelial cells, typhoid toxin interacts with podocalyxin-like protein 1 (PODXL) (Fig. 2a, b), a member of the CD34 sialomucin protein family, which localizes to the apical side of epithelial cells and is also expressed in vascular endothelial cells¹⁵. Consistent with its potential role as a toxin receptor, short hairpin RNA (shRNA)-mediated depletion of PODXL resulted in a significant reduction in toxin binding (Fig. 2c and Supplementary Fig. 3) and toxin-mediated cell-cycle arrest (Fig. 2d). Because we found that typhoid toxin is capable of intoxicating a broad range of cells (Supplementary Fig. 2), we

examined the interaction of typhoid toxin with surface proteins of other cell lines. We found that in macrophages as well as in T and B cells, typhoid toxin interacts with receptor tyrosine phosphatase C, also known as CD45 (Supplementary Fig. 4), which is ubiquitously expressed in haematopoietic cells other than erythrocytes and platelets¹⁶. These results indicate that typhoid toxin may engage different receptors in different cells. The identified typhoid toxin-interacting proteins, however, are all glycosylated, indicating that typhoid toxin may interact with these different surface proteins through common carbohydrate moieties. Consistent with this hypothesis we found that removal of surface glycans significantly reduced typhoid toxin binding to cultured cells (Fig. 2e). Furthermore, a cell line lacking all complex and hybrid N-glycans on glycoproteins due to a mutation in N-acetylglucosaminyltransferase I^{17,18} was more resistant to typhoid toxin than its parent wild-type cell line (Fig. 2f, g and Supplementary Fig. 5). These results indicate that a sugar moiety(s) on surface glycoproteins serves as a receptor for typhoid toxin.

To identify the glycans recognized by typhoid toxin we probed 610 glycans arrayed on a solid surface¹⁹ for binding to biologically active, highly purified, fluorescently labelled typhoid toxin. This analysis revealed a complex binding pattern involving four main glycan groups (Fig. 2h and Supplementary Tables 1 and 2). The first group, which is most commonly present in complex N-linked glycans, is represented by sialylated tri- or bi-antennary glycans (for example, glycans nos 461, 483 and 482) with one or all of the branches terminally sialylated (note: glycan numbers correspond to the nomenclature used by the Consortium for Functional Glycomics, <http://www.functionalglycomics.org>). This group includes glycans with both $\alpha(2-3)$ -linked N-acetylneuraminic acid (Neu5Ac α 2-3; for example, no. 483) and $\alpha(2-6)$ -linked N-acetylneuraminic acid (Neu5Ac α 2-6; for example, no. 482) terminal linkages. However, typhoid toxin probably binds preferentially to Neu5Ac α 2-3 terminal linkages because glycan 483 showed stronger binding than glycan 482, which only differ in their terminal linkages (Fig. 2h and Supplementary Tables 1 and 2). Furthermore, typhoid toxin is likely to bind tri-antennary compounds preferentially over bi-antennary compounds, because glycan 461 showed the highest binding affinity. In addition, the toxin may also bind preferentially the type 2-N-acetylglucosamine unit (Gal β 1-4GlcNAc), present in glycan 461, over type 1-N-acetylglucosamine unit (Gal β 1-3GlcNAc), present in the very similar tri-antennary N-glycan 474, which showed lower binding affinity (Fig. 2h and Supplementary Tables 1 and 2). The second group consists of non-sialylated tri- or bi-antennary glycans also commonly found in complex N-linked glycans. Overall this group had lower binding affinity than sialylated glycans, implying a preference of typhoid toxin for the terminal sialic acid modification. The third group consists of glycans commonly found in glycolipids, mostly as gangliosides. Among these glycans typhoid toxin likely prefers Neu5Ac α 2-8Neu5Ac disialosides, present in ganglioside GD2²⁰ (glycan 228), over Neu5Ac monosialoside, present in GM3 (glycan 263), which did not bind the toxin. These results indicate that, in certain cells, typhoid toxin may also be able to use glycolipids as a receptor. Consistent with this hypothesis, typhoid toxin was able to intoxicate an N-acetylglucosaminyltransferase I-deficient cell line, although only when applied at high concentrations (Supplementary Fig. 6). The fourth group is defined by glycans commonly found on O-glycans, such as glycan 243, which shares the canonical structure of mucin type O-Gal N-acylated glycan. Taken together, these results indicate that typhoid toxin can use as receptors a broad range of glycans, preferentially on surface glycoproteins, but also, although less efficiently, on surface glycolipids, providing a mechanistic explanation for its broad cell target specificity.

To uncover the organization of typhoid toxin, we solved its crystal structure to 2.4 Å resolution. The structure showed a complex of five PltB molecules and one molecule each of PltA and CdtB (Fig. 3a and Supplementary Table 3), which is consistent with the stoichiometry observed by size-exclusion chromatography combined with dynamic light scattering and quantitative amino acid composition analysis (Supplementary Fig. 7). This indicates an unprecedented A₂B₅ organization for typhoid toxin, which is in contrast to other known AB₅

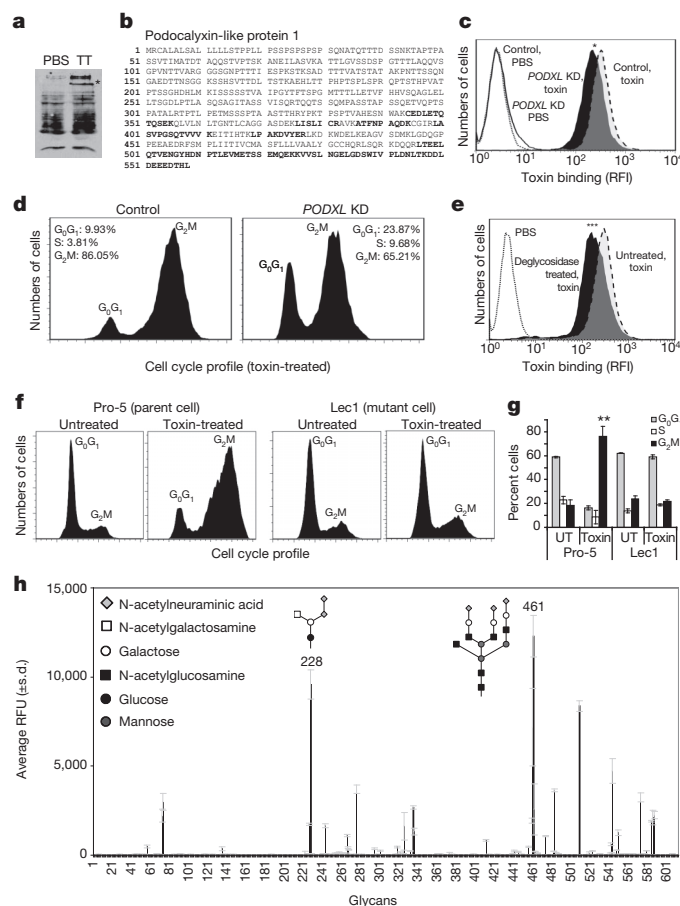


Figure 2 | Typhoid toxin recognizes terminally sialylated glycans on surface glycoproteins. **a**, Henle-407 cell-surface proteins were biotinylated, co-immunoprecipitated with purified typhoid toxin (TT), and analysed by SDS-PAGE and LC-MS/MS. **b**, The peptides from podocalyxin like protein 1 (PODXL) (indicated by an asterisk in **a**) identified by LC-MS/MS are indicated in bold. **c**, PODXL-depleted (by a specifically targeted siRNA) and control cells were treated with fluorescently labelled typhoid toxin and toxin binding was evaluated by flow cytometry (**c**) (* $P = 0.024$ for three independent determinations). KD, knock down. **d**, Alternatively, siRNA-depleted and control cells were treated with typhoid toxin and toxicity was evaluated by cell-cycle analysis. **e**, Henle-407 cells were treated with a mixture of deglycosidases and the ability of treated and control cells to bind fluorescently labelled toxin was subsequently evaluated by flow cytometry (** $P < 0.001$ from three independent experiments). **f**, The N-acetylglucosaminyltransferase I-deficient (Lec1) and its parent (Pro-5) cell lines were treated with typhoid toxin and toxicity was evaluated by cell-cycle analysis. **g**, Quantification of the cell cycle profiles. Bars represent average \pm standard deviation of at least three independent determinations. ** $P < 0.01$, compared to the number of Pro-5 cells in G2M. **h**, Glycan array analysis of typhoid toxin binding. Values represent the average relative fluorescence unit (RFU). The x axis depicts the glycan numbers. The structure of the most relevant glycans is shown.

toxins that have only one A subunit^{9,21}. The pyramid-shaped complex has a height of ~ 90 Å and a maximum width of ~ 60 Å (Fig. 3a), with CdtB located at the vertex, connected by PltA to a pentameric disc at the base of the pyramid made of five monomers of PltB (Fig. 3a, b). The tandem linear arrangement of the enzymatic subunits PltA and CdtB dictates that there are no interactions between CdtB and PltB. As predicted by the amino acid sequence similarities, the PltA and CdtB subunits have a structure very similar to that of the pertussis toxin S1 (and other ADP ribosyl transferases)²² and to the CdtB subunit of cytolethal distending toxin²³. PltA aligns very well to the pertussis toxin S1 domain with a root-mean-squared deviation (r.m.s.d.) of 2.168 Å over 140 C α atoms (with 31% sequence identity) (Supplementary Fig. 8). The positions of the conserved catalytic residues (Glu 133 in typhoid toxin and Glu 129 in pertussis toxin S1), as well as the disulphide bonds (Cys 56–Cys 207 in typhoid toxin and Cys 41–Cys 201 in pertussis toxin S1) overlap almost completely (Supplementary Fig. 8). The latter indicates that, similarly to the pertussis toxin S1 subunit²⁴, PltA would have to be reduced to allow the access of NAD and its putative substrates to the

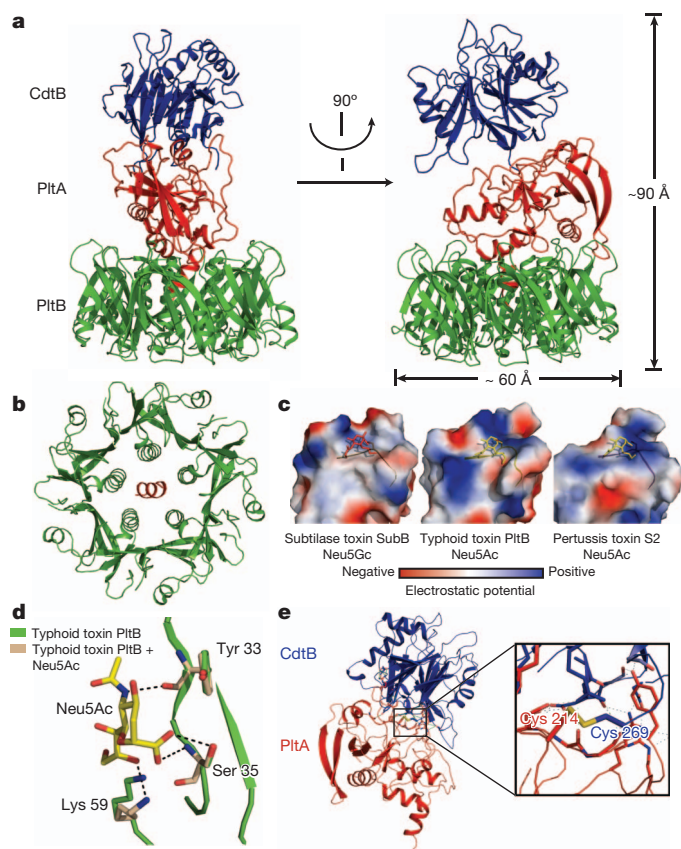


Figure 3 | The crystal structure of typhoid toxin depicts a unique architecture. **a**, Two views of the overall structure of the typhoid holotoxin complex shown as a ribbon cartoon and related by 90° rotation about a vertical axis. CdtB, PltA and PltB are shown in blue, red and green, respectively. **b**, Bottom view of the channel formed by the PltB pentamer (in green), depicting the PltA C-terminal α -helix (in red) within it. **c**, Surface charge distribution of the predicted sugar-binding pockets of different B subunit homologues of the indicated AB₅ toxins (SubB for subtilase and S2 for pertussis toxins). A highly conserved serine residue critical for sugar binding is indicated within the sugar-binding pocket. The sugars N-glycolylneuraminic acid (Neu5Gc; within SubB) and N-acetylneuraminic acid (within typhoid and pertussis toxins) are shown. **d**, Molecular modelling of N-acetylneuraminic acid within the typhoid toxin binding pocket. Critical residues engaged in this interaction are shown. **e**, Atomic interface between CdtB and PltA. The inset shows a detailed view of a critical disulphide bond between PltA Cys 214 and CdtB Cys 269.

active site, and consequently, a reducing activating step must be necessary before contacting its host cell target(s). Typhoid toxin CdtB aligns very well to the *Haemophilus ducreyi* CdtB with an r.m.s.d. of 0.947 Å over 206 C α atoms (with 52% sequence identity) (Supplementary Fig. 9). The positions of the conserved catalytic residues in Typhoid toxin's CdtB overlap almost completely with those of its homologue in *H. ducreyi*.

Similar to the B subunits of other AB₅ toxins⁹, the PltB oligomer is arranged as a pentamer with a central channel that is lined by five helices (Fig. 3b). The PltB protomer shows a typical oligosaccharide-binding fold on the side of the pentamer (Fig. 3c and Supplementary Fig. 10), a location similar to that of toxins that preferentially bind glycoproteins, but different from those that preferentially bind glycolipids, which have the sugar-binding pockets on the membrane-proximal face of the toxin⁹. These findings are consistent with the observation that typhoid toxin preferentially binds glycans present on surface glycoproteins over those present on glycolipids (Supplementary Tables 1 and 2). The monomer of typhoid holotoxin PltB aligns very well with SubB, the subtilase cytotoxin B subunit, with an r.m.s.d. of ~ 0.5 Å over 97 C α atoms (with 50% sequence identity) (Supplementary Fig. 10). Of note, the position of the conserved putative sugar-binding residue Ser 35 overlaps very well (Fig. 3c and Supplementary Fig. 10). The predicted sugar-binding pocket in PltB is not as deep and seems more extended than in SubB, which also differs in surface-charge distribution (Fig. 3c). These differences may account for the significantly different binding specificities shown by these two toxins²⁵. We also compared PltB to the pertussis toxin S2 B subunit²⁶. Although their overall amino sequence similarity is low, the structures can be aligned very well around their sugar-binding domains with an r.m.s.d. of 1.752 Å over 80 C α atoms (Supplementary Fig. 10). Residues in pertussis toxin S2 involved in sugar binding (Tyr 102, Ser 104, Arg 125) are well conserved in PltB (Tyr 33, Ser 35, Lys 59) (Fig. 3d), and the charge distribution and architecture of the sugar-binding pockets are similar (Fig. 3c). This is consistent with the observation that, despite overall less conservation, these two B subunits share sugar-binding specificity. For example, several of the glycans that bind typhoid toxin possess an Neu5Ac moiety at their terminal end, a determinant that also binds pertussis toxin²⁷. Structural modelling of Neu5Ac bound to PltB predicts almost identical interaction to those observed in the atomic structure of pertussis toxin bound to Neu5Ac (Fig. 3c, d and Supplementary Fig. 11). Consistent with the structural predictions, a mutation in the sugar-binding pocket of PltB (PltB^{S35A}) abrogated the ability of typhoid toxin to bind glycans in glycoarray (Fig. 4a and Supplementary Table 4) and surface plasmon resonance assays (Supplementary Fig. 12), and the ability of the toxin to bind (Fig. 4b) and intoxicate (Fig. 4c) cultured cells or to cause symptoms when systemically applied to mice (Fig. 4d). These results are also consistent with the notion that there is a single carbohydrate-binding domain in typhoid toxin because the mutant abrogated binding to all carbohydrates. Surface plasmon resonance assays also indicated that on average, at least for the glycan tested (GD2), each PltB pentamer binds 2.5 sugar molecules with an affinity of ~ 1.2 mM (Supplementary Fig. 12). However, the glycoarray analysis predicts that the binding affinity is likely to be higher for more complex glycans.

The interaction of PltA with the PltB oligomer occurs largely through its carboxy terminus, which buries $1,657$ Å² and has a ΔG of -17.7 kcal mol⁻¹. A critical element in this interaction is a short helix at the carboxy-terminal end of PltA, which inserts into the hydrophobic lumen of the PltB channel (Fig. 3b and Supplementary Fig. 13) and stabilizes the complex by critical interactions mediated by Pro 234, Val 235, Ile 236, Leu 238, Ile 239 and Leu 240 in PltA.

The most salient feature of the interface between PltA and CdtB is a disulphide bond between Cys 214 of PltA and Cys 269 of CdtB (Fig. 3e). The rest of the interface between the two A subunits is unremarkable, burying only 950 Å² with a ΔG of only -5.2 kcal mol⁻¹, indicating that the disulphide bond is essential for the integrity of the holotoxin. Consistent with this hypothesis, subjecting the typhoid toxin to reducing conditions resulted in the dissociation of CdtB from

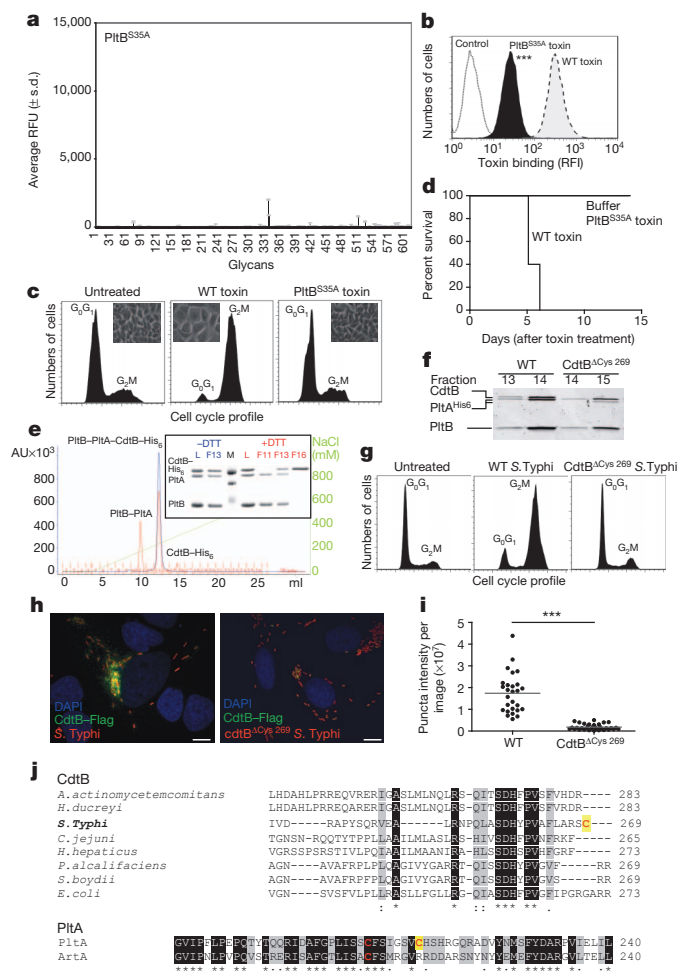


Figure 4 | Structure-function analysis of typhoid toxin. **a, b**, Fluorescently labelled typhoid toxin containing PltB^{S35A} was tested for its binding to glycans (**a**) and to cultured cells (**b**) (see Fig. 2 for details) (***P* < 0.001 from at least three independent determinations). **c, d**, Alternatively, toxicity was assayed by flow cytometric cell-cycle analysis of toxin-treated cells (at least three independent experiments) (**c**), or by systemic administration to mice (*n* = 3 to 5 mice) (**d**). **e**, The typhoid toxin complex was analysed by ion exchange chromatography before (blue) and after (red) treatment with DTT (L, loading control; M, molecular weight markers; F, chromatographic fraction). Inset shows SDS-PAGE analyses of the indicated fractions. **f**, A toxin preparation obtained from a bacterial strain expressing CdtB^{ΔCys269} was analysed by gel-filtration chromatography and compared to wild-type toxin (the experiment was repeated two times). Whereas wild-type holotoxin eluted in fractions 13 and 14, toxin obtained from a bacterial strain encoding CdtB^{ΔCys269} eluted in fractions 14 and 15 due to the lack of CdtB. **g**, Henle-407 cells were infected with *S. Typhi* strains encoding Flag-tagged CdtB or CdtB^{ΔCys269} and cells were examined for toxicity by flow cytometry. **h, i**, Alternatively, cells were fixed, stained with anti-Flag antibody, and the amount of puncta staining, which represent CdtB in typhoid toxin export carriers⁸, were determined by immunofluorescence analysis. Bar represents average of puncta-associated fluorescence intensity (at least 100 cells were analysed in three independent experiments). ****P* < 0.0001. Scale bar, 10 μ m. **j**, ClustalW amino acid sequence comparison of CdtB and PltA homologues. Conserved cysteines are shown in red and unique cysteines are indicated with a yellow highlight. *A. actinomycetemcomitans*, *Aggregatibacter actinomycetemcomitans*; *C. jejuni*, *Campylobacter jejuni*; *H. ducreyi*, *Haemophilus ducreyi*; *H. hepaticus*, *Helicobacter hepaticus*; *P. alcalifaciens*, *Providencia alcalifaciens*; *S. boydii*, *Shigella boydii*.

the PltA–PltB complex (Fig. 4e). Furthermore, introduction of a mutation in Cys 269 of CdtB destabilized the holotoxin complex *in vitro* (Fig. 4f), and resulted in a loss of CdtB-dependent toxicity in *S. Typhi*-infected cells (Fig. 4g and Supplementary Fig. 14). This mutation also

prevented the assembly of CdtB into its vesicle carrier intermediates that can be visualized as fluorescent puncta in *S. Typhi*-infected cells (Fig. 4h, i), despite the fact that the mutant was expressed at equivalent levels to those of wild type (Supplementary Fig. 15). These results indicate that the CdtB^{C269}–PltA^{C214} disulphide bond is critical for the assembly of typhoid toxin holotoxin in the periplasm of the bacteria. Remarkably, even though the sequence conservation between typhoid toxin's CdtB and that of its cytolethal distending toxin homologues is very high, the Cys 269 is unique to Typhoid toxin's CdtB (Fig. 4j). Likewise, although all close homologues of PltA including ArtA, a closely related mono ADP ribosyl transferase from *Salmonella enterica* serovar Typhimurium²⁸, have only two Cys that form an intramolecular disulphide bond, PltA is unique in that it has a third Cys (Cys 214) to pair with *S. Typhi*'s CdtB (Fig. 4j). Therefore typhoid toxin's CdtB has been specifically adapted for its tethering to the PltA–PltB complex by the evolution of uniquely positioned Cys residues in both PltA and CdtB. The covalent linkage of CdtB to the PltA–PltB complex by a disulphide bond ensures that CdtB and PltA are delivered simultaneously to the same target cell. Furthermore, this arrangement would ensure that, after endocytosis and retrograde transport to its place of translocation (most likely the endoplasmic reticulum), PltA and CdtB would be freed from one another upon reduction of the disulphide bond by local reductases, thus allowing them to become competent for translocation to the cytosol and delivery to their place of action.

This study provides unique insight into typhoid toxin, a critical virulence factor of *S. Typhi*, revealing unprecedented features for a bacterial exotoxin and providing significant information on the pathogenesis of typhoid fever. Although human experimentation will be required to confirm the relevance of these findings, the potential implication of typhoid toxin in the development of symptoms during the acute, life-threatening phase of typhoid fever may facilitate the development of potentially life-saving therapeutics against this disease.

METHODS SUMMARY

Wild-type *S. Typhi* strains have been described previously²⁹. Mutants were constructed as described previously³⁰. Protein purification and analysis, crystallization and structure determination were carried out following standard methods. Bacterial infections, mammalian cell intoxication and immunofluorescence microscopy were carried out as described previously⁷. Typhoid toxin binding assays were carried out with Oregon Green-488 (Invitrogen)-labelled toxin preparations by flow cytometry. Glycan array analysis was carried out at the Consortium for Functional Glycomics Protein–Glycan Interaction Core, at Emory University (<http://www.functionalglycomics.org>).

Full Methods and any associated references are available in the online version of the paper.

Received 18 February; accepted 12 June 2013.

Published online 10 July 2013.

- Parry, C. M., Hien, T. T., Dougan, G., White, N. J. & Farrar, J. J. Typhoid fever. *N. Engl. J. Med.* **347**, 1770–1782 (2002).
- Parkhill, J. et al. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413**, 848–852 (2001).
- Butler, T. Treatment of typhoid fever in the 21st century: promises and shortcomings. *Clin. Microbiol. Infect.* **17**, 959–963 (2011).
- Crump, J. A. & Mintz, E. D. Global trends in typhoid and paratyphoid fever. *Clin. Infect. Dis.* **50**, 241–246 (2010).
- Sabbagh, S. C., Forest, C. G., Lepage, C., Leclerc, J. & Daigle, F. So similar, yet so different: uncovering distinctive features in the genomes of *Salmonella enterica* serovars Typhimurium and Typhi. *FEMS Microbiol. Lett.* **305**, 1–13 (2010).
- Haghjoo, E. & Galán, J. E. *Salmonella typhi* encodes a functional cytolethal distending toxin that is delivered into host cells by a bacterial-internalization pathway. *Proc. Natl Acad. Sci. USA* **101**, 4614–4619 (2004).
- Spanò, S., Ugaldé, J. E. & Galán, J. E. Delivery of a *Salmonella Typhi* exotoxin from a host intracellular compartment. *Cell Host Microbe* **3**, 30–38 (2008).
- Spanò, S. & Galán, J. E. A novel pathway for exotoxin delivery by an intracellular pathogen. *Curr. Opin. Microbiol.* **11**, 15–20 (2008).
- Beddoe, T., Paton, A., Le Nours, J., Rossjohn, J. & Paton, J. Structure, biological functions and applications of the AB5 toxins. *Trends Biochem. Sci.* **35**, 411–418 (2010).

10. Lara-Tejero, M. & Galán, J. E. A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. *Science* **290**, 354–357 (2000).
11. Lara-Tejero, M. & Galán, J. E. Cytolethal distending toxin: limited damage as a strategy to modulate cellular functions. *Trends Microbiol.* **10**, 147–152 (2002).
12. Charles, R. C. *et al.* Characterization of anti-*Salmonella enterica* serotype Typhi antibody responses in bacteremic Bangladeshi patients by an immunoaffinity proteomics-based technology. *Clin. Vaccine Immunol.* (2010).
13. Liang, L. *et al.* Immune profiling with a *Salmonella* Typhi antigen microarray identifies new diagnostic biomarkers of human typhoid. *Scientific Rep.* **3**, 1043 (2013).
14. Connor, B. A. & Schwartz, E. Typhoid and paratyphoid fever in travellers. *Lancet Infect. Dis.* **5**, 623–628 (2005).
15. Yu, C.-Y. *et al.* A bipartite signal regulates the faithful delivery of apical domain marker podocalyxin/Gp135. *Mol. Biol. Cell* **18**, 1710–1722 (2007).
16. Hermiston, M. L., Zikherman, J. & Zhu, J. W. CD45, CD148, and Lyp/Pep: critical phosphatases regulating Src family kinase signaling networks in immune cells. *Immunol. Rev.* **228**, 288–311 (2009).
17. Kumar, R., Yang, J., Larsen, R. & Stanley, P. Cloning and expression of N-acetylglucosaminyltransferase I, the medial Golgi transferase that initiates complex N-linked carbohydrate formation. *Proc. Natl Acad. Sci. USA* **87**, 9948–9952 (1990).
18. Stanley, P., Narasimhan, S., Siminovitch, L. & Schachter, H. Chinese hamster ovary cells selected for resistance to the cytotoxicity of phytohemagglutinin are deficient in a UDP-N-acetylglucosamine-glycoprotein N-acetylglucosaminyltransferase activity. *Proc. Natl Acad. Sci. USA* **72**, 3323–3327 (1975).
19. Song, X. *et al.* Shotgun glycomics: a microarray strategy for functional glycomics. *Nature Methods* **8**, 85–90 (2011).
20. Yu, R. K., Tsai, Y.-T., Ariga, T. & Yanagisawa, M. Structures, biosynthesis, and functions of gangliosides—an overview. *J. Oleo Sci.* **60**, 537–544 (2011).
21. Merritt, E. A. & Hol, W. G. J. AB₅ toxins. *Curr. Opin. Struct. Biol.* **5**, 165–171 (1995).
22. Stein, P. E. *et al.* The crystal structure of pertussis toxin. *Structure* **2**, 45–57 (1994).
23. Nešić, D., Hsu, Y. & Stebbins, C. E. Assembly and function of a bacterial genotoxin. *Nature* **429**, 429–433 (2004).
24. Locht, C., Coutte, L. & Mielcarek, N. The ins and outs of pertussis toxin. *FEBS J.* **278**, 4668–4682 (2011).
25. Byres, E. *et al.* Incorporation of a non-human glycan mediates human susceptibility to a bacterial toxin. *Nature* **456**, 648–652 (2008).
26. Stein, P. E. *et al.* Structure of a pertussis toxin-sugar complex as a model for receptor binding. *Nature Struct. Biol.* **1**, 591–596 (1994).
27. Millen, S. H., Lewallen, D. M., Herr, A. B., Iyer, S. S. & Weiss, A. A. Identification and characterization of the carbohydrate ligands recognized by pertussis toxin via a glycan microarray and surface plasmon resonance. *Biochemistry* **49**, 5954–5967 (2010).
28. Saitoh, M. *et al.* The *artAB* genes encode a putative ADP-ribosyltransferase toxin homologue associated with *Salmonella enterica* serovar Typhimurium DT104. *Microbiology* **151**, 3089–3096 (2005).
29. Galán, J. E. & Curtiss, R. III. Distribution of the *invA*, *-B*, *-C*, and *-D* genes of *Salmonella typhimurium* among other *Salmonella* serovars: *invA* mutants of *Salmonella typhi* are deficient for entry into mammalian cells. *Infect. Immun.* **59**, 2901–2908 (1991).
30. Kaniga, K., Bossio, J. C. & Galán, J. E. The *Salmonella typhimurium* invasion genes *invF* and *invG* encode homologues of the AraC and PulD family of proteins. *Mol. Microbiol.* **13**, 555–568 (1994).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank members of the Galán laboratory for careful review of this manuscript. We thank E. Foltá-Stogniew for help with conducting the surface plasmon resonance and light-scattering size-exclusion chromatography assays. We also thank M. Lara-Tejero and X. Liu for mass spectrometry analysis, J. M. Kim for help with glycan array analysis software, and K.-W. Kim for help in animal inoculations. We thank W. Meng for providing help with X-ray diffraction data collection, J. Wang and C. Yan for suggestions and providing help with structure refinement, and X. Gong and M. Ke for help and suggestions with molecular docking. The glycan array analysis was carried out at the Consortium for Functional Glycomics Protein-Glycan Interaction Core, at Emory University, which is supported by PHS Grant GM098791. J.S. was supported in part by a grant from the Northeast Biodefense Center U54-AI057158 and this work was supported by NIAID Grant AI079022 to J.E.G.

Author Contributions J.S., X.G. and J.E.G. designed the studies and interpreted the results. J.S. and X.G. carried out the experiments. J.S., X.G. and J.E.G. prepared the manuscript.

Author Information The atomic coordinates have been deposited in the RCSB Protein Data Bank (entry number 4K6L). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.E.G. (jorge.galan@yale.edu).

METHODS

Bacterial strains, mammalian cells and growth conditions. The wild-type *Salmonella enterica* serovar Typhi strain ISP2825 has been described previously²⁹. A derivative of this strain encoding a Flag-epitope tagged CdtB has been previously described⁷. Other *S. Typhi* mutant strains were constructed by allelic exchange following previously described procedures³⁰. Site-directed mutations and epitope tagging was carried out following standard recombinant DNA procedures. For all infection experiments, *S. Typhi* strains were grown at 37 °C in 2 ml LB broth containing 0.3 M NaCl to an OD_{600 nm} of ~0.9 after inoculation from overnight cultures at a dilution of 1:50. Cultured cells and culture medium used in these studies were as follows. Henle-407 (human intestinal epithelial cells): DMEM + 10% BCS; Jurkat (human T lymphocytes): RPMI1640 + 10% FBS; Ramos (human B lymphocytes): RPMI1640 + 10% FBS; THP1 (human monocytic cells): RPMI1640 + 10% FBS + 0.05 mM β -mercaptoethanol; Raw (mouse monocytes/macrophages): DMEM + 10% FBS; NIH3T3 (mouse embryonic fibroblasts): DMEM + 10% BCS; COS1 (monkey kidney fibroblasts): DMEM + 10% BCS; CHO (Chinese hamster ovary epithelial cells): Ham's F12 + 10% FBS; MDCK (canine kidney epithelial cells): MEM + 10% FBS; Lec1 (N-acetylglucosaminyltransferase I mutant) and parent cell Pro-5: alpha MEM with ribonucleosides and deoxyribonucleosides + 10% FBS. All mammalian cells were originally obtained from American Type Culture Collection and cultured at 37 °C under an atmosphere of 5% CO₂.

Typhoid toxin expression and purification. *plbB*, *plbA* and *cdtB* (wild-type or mutant alleles, and 3 \times Flag- or 6 \times His-tagged at the carboxy terminus of CdtB or internally in the case of PltA, as indicated) were cloned as a single operon in either a low copy plasmid derived from pWSK129³¹, or in pET28a+ (Novagen). *Escherichia coli* strains carrying the different plasmids were grown to an OD_{600 nm} of 0.6–0.7 at 37 °C; expression of typhoid toxin was subsequently induced by the addition of 0.5 mM IPTG, and induced cultures were incubated overnight at 30 °C. Bacterial cells were pelleted by centrifugation, and bacterial cells were resuspended in a buffer containing 200 mM Tris-HCl (pH 7.5), 20% sucrose, 1 mM EDTA, 2 mg ml⁻¹ lysozyme, and 1 \times protease inhibitors (2 ml per each gram of wet cell pellet weight), incubated for 5 min at room temperature, mixed with dH₂O (3 ml per each gram of wet cell pellet weight) by inversion, and incubated for an additional 10 min on ice. A crude periplasmic protein fraction was obtained by centrifugation at 4,000g for 15 min at room temperature, and used as a source of typhoid toxin for affinity chromatography purification using M2 Flag affinity gel (Sigma). The periplasmic fraction containing typhoid holotoxin was incubated for 3 h at 4 °C with M2 agarose beads packed on a 10 ml column (Bio-Rad), washed with ~20 bed volume of PBS, and eluted twice with PBS containing 3 \times Flag peptide (Sigma; 150 ng μ l⁻¹). Partially purified holotoxin was applied on a Superdex 200 size-exclusion chromatography (Tris-HCl buffer, 15–50 mM, in a pH range of 7.6–8.0 supplemented with 150 mM NaCl) to complete its purification. His-epitope-tagged typhoid toxin was purified as follows. *E. coli* cultures were resuspended in a buffer containing 15 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mg ml⁻¹ lysozyme, 10 μ g ml⁻¹ DNase and 1 \times protease inhibitor cocktail, lysed by passing them through a French press three times, pelleted, and affinity-purified using a Nickel-resin (Qiagen) according to the vendor's recommendation. The eluates were diluted in 20 mM MES, pH 6.0 buffer and loaded onto a Hi Trap ion-exchange column. Fractions from the ion-exchange chromatography were monitored on SDS-PAGE, concentrated, and further purified by using a Superdex 200 column. Final fractions were examined for purity on a 15% SDS-PAGE.

Mammalian cell intoxication assay. Cell-cycle arrest after typhoid toxin intoxication was examined by flow cytometry as previously described⁷. Briefly, after treatment with 3 \times Flag-tagged typhoid toxin or bacterial infection for different times (as indicated), cells were trypsinized, collected, washed and fixed overnight at -20 °C in ~70% ethanol/PBS. Fixed cells were washed with PBS and resuspended in 500 μ l of PBS containing 50 μ g ml⁻¹ propidium iodide, 0.1 mg ml⁻¹ RNase A and 0.05% Triton X-100. After incubation for 40 min at 37 °C, cells were washed with PBS, resuspended in 500 μ l PBS, filtered and analysed by a flow cytometry. DNA contents of cells was determined using FlowJo (Treestar).

Light-scattering size exclusion chromatography and amino acid composition analysis. Light-scattering size exclusion chromatography (SEC-LS) and amino acid composition analysis were carried out at the Keck Biotechnology Resource Laboratory at the Yale University School of Medicine (<http://medicine.yale.edu/keck/index.aspx>). For SEC-LS analysis, the toxin was purified in 50 mM HEPES buffer containing 150 mM NaCl and run on a Superdex 200 column equipped with a light scattering detector using a same HEPES buffer. For amino acid composition analysis, typhoid toxin was resolved on a 15% SDS-PAGE gel, stained with Coomassie brilliant blue, and the three individual bands were excised and used for amino acid composition analysis.

Mouse intoxication experiments. All animal experiments were conducted according to protocols approved by Yale University's Institutional Animal Care and Use Committee. Age- and sex-matched 7–14 weeks old C57BL/6 mice were

randomly allocated in each group and numbers of animals per each group were estimated to obtain statistically significant data after a pilot experiment to get a sense of the potential variability. Groups of mice were intravenously injected with 100 μ l solutions containing either TBS buffer alone or 10 μ g of each of the purified holotoxin preparations (endotoxin free). His-tagged wild-type, PltA catalytic mutant (PltA^{E133A}), CdtB catalytic mutant (CdtB^{H160Q}), double catalytic mutant (PltA^{E133A}CdtB^{H160Q}), and PltB^{S35A} mutant holotoxins were purified as indicated above. Changes in behaviour, weight and temperature of the toxin-injected mice as well as their survival were closely monitored during the duration of the experiment. To minimize bias, blind end-point assessment was applied to examine the behaviour of the animals after toxin inoculation.

Peripheral blood leukocyte preparation, immunostaining and flow cytometry analysis. Peripheral blood samples of typhoid toxin treated and control mice were collected into heparinized tubes, incubated with 1 ml ACK lysis buffer (BioWhittaker) (to remove red blood cells), incubated for 5 min on ice, washed with 2 ml PBS, and centrifuged to collect peripheral blood leukocytes (PBLs). After a repetition of the red blood cell removal step, PBLs were used for immunostaining as described below. After washing, the cells were immediately incubated for 30 min on ice with 100 μ l of anti-mouse Ly-6G (Gr-1) antibody conjugated with FITC (eBioscience, cat. no. 11-5931-81). The cells were then washed with 2 ml FACS buffer (PBS, 0.16% BSA), resuspended in 100 μ l FACS fixation buffer (PBS, 1% paraformaldehyde, 1% FCS), and used for flow cytometric analyses on a FACScalibur flow cytometer (BD Biosciences). Alternatively, blood samples collected by heart puncture 4 days after toxin treatment were analysed in a Hemavet 950FS haematology analyser (Drew Scientific).

Identification of typhoid toxin interacting, biotin-labelled host cell surface proteins. Cultured cells (Henle, Jurkat, Ramos or THP1 at ~50% confluency) were washed with PBS, treated with PBS containing ~100 μ g ml⁻¹ of Sulfo-NHS-SS-Biotin (Thermo) for 30 min at room temperature, and subsequently washed 3 times with 50 mM Tris-HCl pH 8.0/150 mM NaCl buffer to quench and to remove extra biotin reagent. After additional washings (3 times with PBS), cells were resuspended in a lysis buffer containing 50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% NP-40, and 1 \times protease inhibitor, incubated for 20 min on ice, and homogenized by passing a 26-G needle ~20 times. After removal of cellular debris by centrifugation, the supernatants were mixed with 10 μ g purified Flag-tagged typhoid toxin and incubated for 3 h at 4 °C. Anti-Flag antibody-containing agarose beads were added, incubated for additional 1 h at 4 °C, washed 5 times with PBS, SDS-PAGE sample buffer was added, boiled, and run on SDS-PAGE. The gels were transferred to nitrocellulose membranes, blocked with TBS containing 5% BSA, incubated overnight at 4 °C with streptavidin conjugated to horseradish peroxidase (1:5,000) in TBS with 1% BSA, washed with TBST, and developed with ECL substrates (Pierce). To identify the typhoid toxin-interacting proteins by LC-MS/MS, equivalently obtained samples were run in parallel, stained with Coomassie blue, and gel regions corresponding to the molecular weight of typhoid-toxin interacting proteins identified by western blot analysis were excised and processed for LC-MS/MS as previously described³². Briefly, gel slices were destained in destaining buffer (50 mM NH₄HCO₃, 50% acetonitrile (ACN)), and dehydrated with ACN. Disulphide bonds were reduced by incubating the samples with NH₄HCO₃ containing 10 mM DTT and alkylated by incubating them with 55 mM 2-iodoacetamide in 100 mM NH₄HCO₃ buffer for 20 min at room temperature. Gel pieces were dehydrated, trypsin-digested overnight, extracted, run on an LTQ-Velos Mass Spectrometer, and spectra analysed with Mascot (Matrixscience). As negative controls, equivalently processed samples obtained with two irrelevant baits (glutathione S-transferase-conjugated with 3 \times Flag and InvC-3 \times Flag) were used.

Oregon Green 488 typhoid toxin labelling. Purified wild type and PltB^{S35A} mutant typhoid toxin preparations were fluorescently labelled with Oregon Green (OG)-488 dye (Invitrogen) according to the vendor's recommendation. OG-488 dye has a succinimidyl ester moiety that reacts with primary amines of proteins to form stable dye-protein conjugates. Purified toxin preparations (1 mg ml⁻¹) were incubated with reactive dye in 500 μ l of 100 mM bicarbonate buffer for 1 h at room temperature, and applied to a size-exclusion chromatography column provided by the vendor to separate the dye-protein conjugates from free dye. Degree of labelling was determined by measuring the absorbance of the conjugate solution at 280 nm and 496 nm, which yielded comparable toxin labelling for both toxin preparations (4.4:1 and 4.36:1 dye/holotoxin ratios for wild-type and PltB^{S35A} mutant toxin, respectively). A predicted extinction coefficient of 191,400 M⁻¹ cm⁻¹ was used to calculate the dye/toxin ratio.

Typhoid toxin binding assay. Cells were collected by trypsinization, washed with HBSS, resuspended in 100 μ l HBSS containing 0.2 μ g of Oregon Green-488 (Invitrogen)-labelled purified wild type or mutant toxin preparations. Cells were incubated in the presence of the labelled toxin preparations for 30 min at room temperature, washed with PBS, resuspended in 100 μ l PBS containing 1% paraformaldehyde, and

analysed by flow cytometry. When indicated, Henle-407 cells were treated for 2 h with 10 μ l deglycosidase mix (NEB) in 2 ml HBSS before processing for toxin-binding assays as indicated above.

Amino acid sequence alignment. ClustalW amino acid sequence comparison analyses of CdtB and PltA homologues was carried out using the following sequences (GenBank entry numbers): CdtB homologues from *Shigella boydii* (AAU88264.1), *Providencia alcalifaciens* (BAL72684.1), *Helicobacter hepaticus* (AAF19158.1), *Haemophilus ducreyi* (NP_873398.1), *E. coli* (BAH72965.1), *Campylobacter jejuni* (AAS01598.1), and *Aggregatibacter actinomycetemcomitans* (AAC70898.1). The PltA homologue used in the alignment was the highly related *S. Typhimurium* DT104 ArtA (BAE20153.1).

Generation of PODXL-depleted cell lines. RNA interference vector pSUPER-H1 (Oligoengine) was used to generate a plasmid expressing an shRNA construct targeted to PODXL. Oligonucleotides including a target region for PODXL 5'-GATCCCGGACAAATGGGATGAAGTATTCAAGAGATAGTTCATCCCA-TTTGTCTTTTTC and 5'-TCGAGAAAAAGGACAAATGGGATGAAGTATCTCTTGAATAGTTCATCCCATTTGTCCGGG were annealed to form double-stranded DNA and cloned into the BglII and HindIII sites of pSUPER-H1 vector. Henle-407 cells were transfected with this plasmid using Lipofectamine 2000 (Invitrogen) and puromycin-resistant stable-transfected cell lines were screened for PODXL expression by real time-PCR using a PODXL-specific primer set. The primer sequences were as follows: 5'-ACCGGGGACTACAACCCTG (sense) and 5'-TGTTGGTGTAGGTTTAGCTGTG (antisense) for PODXL and 5'-GATTA CTGCTCTGGCTCCTAGC (sense) and 5'-GACTCATCGTACTCTGCTTGC (antisense) for β -actin.

Glycan array analysis. OG488-labelled wild type and PltB^{S35A} mutant holotoxins were diluted to 180 μ g ml⁻¹ or 20 μ g ml⁻¹ and an aliquot (70 μ l) was applied to separate microarray slides (version 5.1) at the Consortium for Functional Glycomics Protein-Glycan Interaction Core, at Emory University (<http://www.functionalglycomics.org>). The data are reported as average relative fluorescence units of four of six replicates (after removal of the highest and lowest values) for each glycan represented on the array. Glycans showing typhoid toxin binding activity (listed in Supplementary Table 1) were selected considering a cut off value that was larger or equal to than 1% of the values obtained with the glycan showing the highest binding activity. Glycans showing a variation coefficient higher than 30% were eliminated from this group. In addition some specific glycans were eliminated from the group because they are physiologically irrelevant (glycan 509) or showed non-specific binding (glycans 335, 336, 523).

Surface plasmon resonance. Surface plasmon resonance analysis was carried out at the Keck Biotechnology Resource Laboratory at the Yale University School of Medicine (<http://medicine.yale.edu/keck/index.aspx>) using a BiaCore biosensor. Briefly, 50 μ g ml⁻¹ anti-M2 Flag antibody was immobilized on the surface of a chip by amine coupling. Purified wild-type or PltB^{S35A} (both Flag-tagged on the C terminus of CdtB) mutant toxin preparations were applied to the chip followed by application of ganglioside GD2 glycan (Elicityl) at various concentrations.

Crystallization. The purification of 6 \times His-tagged typhoid toxin used for crystallization is described above. Initial sparse matrix crystallization trials of full-length holotoxin protein preparations (2 mg ml⁻¹) were carried out at the Yale University School of Medicine Structural Biology Core facility. After crystal optimization trials, full length typhoid toxin (4.5 mg ml⁻¹) crystals grew in \sim 3 weeks at room temperature using the hanging-drop vapour-diffusion method in a mix of 1 μ l of protein with 1 μ l of reservoir solution consisting of 1.6 M sodium formate and 0.1 M sodium acetate, pH 4.5.

X-ray data collection and structure determination. X-ray data were collected to 2.4 Å at the wavelength of 1.5418 Å on a Rigaku Homelab system at the Yale University Chemical and Biophysical Instrumentation Center (CBIC) (<http://cbic.yale.edu>). Data were integrated and scaled using the HKL-2000 package³³. Further processing was performed with programs from the CCP4 suite³⁴. The holotoxin structure was determined by molecular replacement using PHASER³⁵ with the

atomic coordinates of chain B of *H. ducreyi* CDT (PDB ID 1SR4)²³, chain A of pertussis toxin (PDB ID 1PRT)²⁶ and subtilase cytotoxin B-subunit (PDB ID 3DWP)²⁵ as the initial search model. To complete the model, manual building was carried out in COOT³⁶. Figures were prepared using PyMol³⁷. The structure refinement was done by PHENIX³⁸. Ramachandran plot statistics (%): Most favourable: 90.1; additionally allowed: 8.8; generously allowed: 1.1; disallowed: 0.0. The data collection and refinement statistics are summarized in Supplementary Table 2.

Molecular docking. Molecular docking of Neu5Ac onto PltB was carried out with AutoDock Vina³⁹. Based on the available structural and functional information on pertussis and subtilase toxins as well as our own functional data indicating the importance of PltB^{Ser35} in sugar binding (see Fig. 4 and Supplementary data), we chose to model the binding of Neu5Ac onto a pocket surrounding PltB Ser 35 and consider several amino acid residues (Tyr 33, Tyr 34, Ser 35 and Lys 59) as flexible. The calculation yielded 20 possible models, of which the one with the highest ranking was selected as the most likely.

Fluorescence microscopy. Fluorescence microscopy analysis of typhoid toxin in *S. Typhi*-infected cells was carried out as previously described⁷. Briefly, Henle-407 cells were seeded on coverslips placed within 24-well plates and cultured overnight. Cultured cells were infected with different strains of *S. Typhi* expressing Flag-epitope tagged typhoid toxin with a multiplicity of infection of 20 for 1 h. Infected cells were washed, treated for 1 h with 100 μ g ml⁻¹ gentamicin to kill extracellular bacteria, washed again, and incubated for 24 h in a cell culture medium containing 10 μ g ml⁻¹ gentamicin. Infected cells were washed with PBS, fixed with 4% paraformaldehyde for 10 min at room temperature, washed and incubated for 30 min at room temperature with PBS containing 50 mM NH₄Cl, 0.2% Triton X-100 and 3% BSA. Cells were then incubated for 30 min at room temperature with a primary antibody mixture of mouse anti-Flag M2 (Sigma; 1:4,000; catalogue no. F3165) and rabbit anti-*S. Typhi* (Difco; 1:4,000; catalogue no. 240993) in PBS containing 3% BSA, washed, incubated for 30 min at room temperature with a secondary antibody mixture of Alexa-488 anti-mouse (1:2,000; catalogue no. A21121) and Alexa-594 anti-rabbit (1:2,000; catalogue no. A11012) in PBS containing 3% BSA. Cells were washed, stained with DAPI (1:10,000), washed again, mounted and viewed using a \times 100 objective on a fluorescence microscope (Nikon TE2000). Puncta intensities of images were analysed using ImageJ software as previously described⁴⁰.

Statistics. The two-tailed student *t*-test was performed to determine the statistical significance of experimental changes from control values. A *P* value of less than 0.05 was considered significant.

- Wang, R. F. & Kushner, S. R. Construction of versatile low-copy-number vectors for cloning, sequencing and gene expression in *Escherichia coli*. *Gene* **100**, 195–199 (1991).
- Liu, X., Gao, B., Novik, V. & Galán, J. E. Quantitative proteomics of intracellular *Campylobacter jejuni* reveals metabolic reprogramming. *PLoS Pathog.* **8**, e1002562 (2012).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computational Project, number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
- Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
- DeLano, W. L. S. The PyMOL Molecular Graphics System. <http://www.pymol.org> (2002).
- Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
- Spanò, S., Liu, X. & Galán, J. E. Proteolytic targeting of Rab29 by an effector protein distinguishes the intracellular compartments of human-adapted and broad-host *Salmonella*. *Proc. Natl Acad. Sci. USA* **108**, 18418–18423 (2011).

Three-state mechanism couples ligand and temperature sensing in riboswitches

Anke Reining¹, Senada Nozinovic¹, Kai Schlepckow¹, Florian Buhr¹, Boris Fürtig¹ & Harald Schwalbe¹

Riboswitches are *cis*-acting gene-regulatory RNA elements that can function at the level of transcription, translation and RNA cleavage^{1–3}. The commonly accepted molecular mechanism for riboswitch function proposes a ligand-dependent conformational switch between two mutually exclusive states⁴. According to this mechanism, ligand binding to an aptamer domain induces an allosteric conformational switch of an expression platform, leading to activation or repression of ligand-related gene expression⁵. However, many riboswitch properties cannot be explained by a pure two-state mechanism. Here we show that the regulation mechanism of the adenine-sensing riboswitch, encoded by the *add* gene on chromosome II of the human Gram-negative pathogenic bacterium *Vibrio vulnificus*⁶, is notably different from a two-state switch mechanism in that it involves three distinct stable conformations. We characterized the temperature and Mg^{2+} dependence of the population ratios of the three conformations and the kinetics of their interconversion at nucleotide resolution. The observed temperature dependence of a pre-equilibrium involving two structurally distinct ligand-free conformations of the *add* riboswitch conferred efficient regulation over a physiologically relevant temperature range. Such robust switching is a key requirement for gene regulation in bacteria that have to adapt to environments with varying temperatures. The translational adenine-sensing riboswitch represents the first example, to our knowledge, of a temperature-compensated regulatory RNA element.

The *add* gene contains a structured 112-nucleotide long sequence (Supplementary Fig. 1) within the 119-nucleotide long 5'-untranslated region (5'-UTR) that acts as an adenine-sensing, translational on-riboswitch. Previous studies have focused on the investigation of adenine aptamer domains alone^{6–11}, and the detailed molecular mechanisms of the conformational switch of full-length riboswitches have remained elusive. The commonly accepted model proposes a ligand-induced switch between two conformations. In any two-state mechanism, complex formation is inherently favoured at low temperatures. Investigating a full-length translational riboswitch^{12–14}, we found a novel riboswitch mechanism involving three conformations. Efficient switching is maintained over a large temperature range, which is important as the pathogen has to replicate both at 37 °C in the human host and at temperatures around 10 °C in its marine habitat.

In the presence of equimolar amounts of adenine, the ligand-bound holo conformation showed unique tertiary interactions (Fig. 1a and Supplementary Figs 2 and 3). The mode of molecular recognition determined previously for the adenine-bound aptamer domain^{15–17} was retained in the full-length riboswitch, as evidenced by the unique pattern of six characteristic NMR imino signals in the binding pocket (Fig. 1b). The three helices P1, P2 and P3 were fully formed and adopted a three-way junction stabilized by long-range interactions between loops L2 and L3. Stabilization of P1 by adenine leads to an opening of P4. Consequently, nucleotides from A111 to U125 were single-stranded, making the Shine–Dalgarno sequence and the AUG start codon accessible to ribosome binding (Fig. 1 and Supplementary

Table 1), in line with the biologically observed activation of gene expression of adenosine deaminase in the presence of adenine¹². This holo conformation of the riboswitch represented the adenine-induced on-state.

Unexpectedly, two ligand-free (apo) conformations instead of one could be detected by NMR. The apoA and apoB conformations revealed markedly different secondary structures (Fig. 1a, Supplementary Fig. 3 and Supplementary Table 2). ApoA was similar to the holo conformation except for the ligand-binding core, the loop–loop interactions and P4. In contrast to previously proposed secondary structure models^{12–14,18}, P1 and P4 were formed, but only partially. Owing to simultaneous partial formation of both helices, only the Shine–Dalgarno, but not the AUG start codon, was masked in the apoA off-state (Fig. 1a).

The apoB conformation differed considerably from both apoA and holo conformations. The main difference was found in the 5' region, in which nucleotides in P1 and P2 formed an alternative secondary structure, a helix–bulge–helix (5 base pairs (bp)–7 nucleotides–2 bp) motif capped by a loop of nine nucleotides. P3, P4 and P5 were present, and formation of the ligand-binding three-way junction was suppressed. Structurally, apoB is binding-incompetent; addition of threefold excess of ligand to the riboswitch did not induce any chemical shift perturbations for apoB.

The three conformations were connected by two equilibria: the binding-competent conformation apoA exchanged with the binding-incompetent conformation apoB (equilibrium ratio between the apo states (K_{pre}) = [apoA]/[apoB]), as monitored by NMR exchange spectroscopy¹⁹. In an adenine-dependent step, apoA could further switch to the holo conformation (dissociation constant (K_d)) (Fig. 1a). The two equilibria were differentially modulated by changes in temperature and Mg^{2+} concentration. K_{pre} changed the population of apoA from 12% to 40% between 10 °C and 30 °C (Fig. 1c). This population ratio was not markedly influenced by Mg^{2+} concentration from 0 to 12 mM ([RNA] = 0.4 mM). By contrast, the binding equilibrium was not only temperature dependent, but also Mg^{2+} dependent. Mg^{2+} stabilized the holo conformation at all temperatures. For example, the addition of 12 mM Mg^{2+} increased the population of holo at 10 °C from 41% to 80%, and at 30 °C from 8% to 76% (Fig. 1c). The stabilizing effect of Mg^{2+} was also reflected in K_d values at different [Mg^{2+}]:[RNA] ratios measured by isothermal titration calorimetry (ITC) (Fig. 1d, e), stopped-flow fluorescence, and NMR (Fig. 2). At 10 °C, the K_d decreased by four orders of magnitude after changing the [Mg^{2+}]:[RNA] ratio from 0 to 2,000:1 (Supplementary Fig. 10). The temperature influence was most prominent at the highest [Mg^{2+}]:[RNA] ratios, for which the K_d increased 100-fold for $\Delta T = 20$ °C.

The rates of interconversion between apoA and apoB were $k_{AB} = 0.91$ s^{–1} and $k_{BA} = 0.40$ s^{–1} at 25 °C (Fig. 1e and Supplementary Fig. 6a, b). In line with RNA transition-state stabilization, increasing the Mg^{2+} concentrations accelerated these rates (Supplementary Fig. 6c), whereas the populations remained largely unaffected. The 5'-terminal first 35 nucleotides of the riboswitch were responsible

¹Center for Biomolecular Magnetic Resonance, Institute of Organic Chemistry and Chemical Biology, Johann Wolfgang Goethe-Universität Frankfurt am Main, Max-von-Laue-Strasse 7, 60438 Frankfurt am Main, Germany.

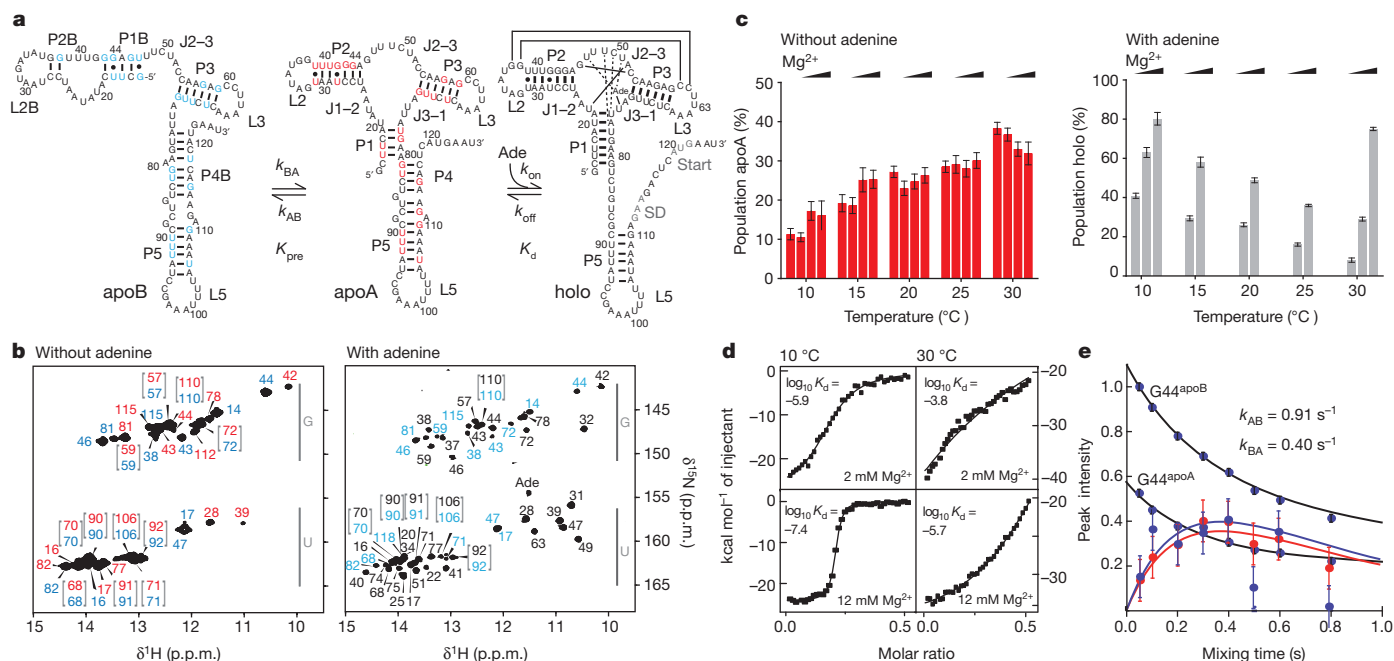


Figure 1 | Modulation of the conformational equilibrium of the wild-type adenine-sensing riboswitch by temperature, Mg^{2+} and cognate ligand.

a, Secondary structure models of the RNA conformations apoA (assigned nucleotides are red), apoB (blue) and holo (black). The constants (K_{pre} and K_{d}) and the rates describing the kinetics of interconversion between the conformations in the two equilibria (k_{AB} , k_{BA} , k_{on} and k_{off}) are depicted. **b**, ^1H - ^{15}N HSQC (heteronuclear single quantum coherence) spectra of the imino group region reporting on the base-pairing pattern of each nucleotide of the wild-type RNA (0.6 mM), Mg^{2+} (5 mM), without adenine (left) and with 1.1 eq adenine (right) recorded at 10 °C and 950 MHz. Assigned peaks are annotated for apoA (red), apoB (blue) and holo (black). Conformation apoA was less than 10% populated in the spectrum with adenine and not visible at the given plot level (Supplementary Fig. 5). **c**, Temperature and Mg^{2+} dependence of the populations of the three wild-type RNA conformations in the absence (left) and presence (right) of adenine, obtained by integrated signals (G115^{apoA}, G44^{apoB}, U47^{holo}) in ^1H - ^{15}N HSQC spectra. Populations of [apoA] = 1 - [apoB] (%), and of [holo] = 1 - [apoA + apoB] (%), are shown.

for the conformational heterogeneity in the apo state. This finding was corroborated by analysis of a 35-nucleotide RNA fragment (G14-U49) that showed the same conformational heterogeneity with identical base pairing in the two underlying secondary structure elements (Supplementary Fig. 7). The recently proposed tuning capability of the P2 region in purine-sensing aptamers is in line with the intrinsic bistable character of the first 35 nucleotides described here²⁰.

To suppress the conformational pre-equilibrium caused by the bistable character and to stabilize the binding-competent conformation apoA, we mutated residues within P2 (Mut^{P2}) (Fig. 2a). The exchange of two AU to GC base pairs (A29C-U41G and A30G-U40C) in P2 (Fig. 2a) led to a thermal stabilization of $\Delta\Delta G = 0.11 \text{ kcal mol}^{-1}$ at 25 °C (Supplementary Fig. 8). In the ligand-free state of Mut^{P2}, a single homogeneous set of NMR signals could be detected (Fig. 2b). P2, P3 and P5 were fully formed, as in the holo conformation of the wild-type sequence (Fig. 2b). In addition, P1 and the interaction between L2 and L3 were already pre-formed in Mut^{P2} and the 3' terminal nucleotides were accessible. No signals reporting the presence of an apoB-like conformation could be detected.

We compared kinetics and thermodynamics of ligand binding to wild-type and Mut^{P2} to determine the effect of the pre-equilibrium at $[\text{Mg}^{2+}]:[\text{RNA}] > 50:1$. Binding of the fluorescent adenine analogue 2-aminopurine was characterized by stopped-flow fluorescence spectroscopy (Fig. 2c, d, Supplementary Fig. 10 and Supplementary Tables 3 and 4), as previously conducted⁷. For Mut^{P2}, the K_{d} values were

significantly lower (by a factor of 30 at 10 °C) than the apparent K_{d} values for the wild type. This discrepancy between the wild-type and Mut^{P2} data can, however, be reconciled if the absolute RNA concentration is corrected for the concentration of the binding-incompetent apoB conformation (Supplementary Fig. 10). Then, ligand affinity to either apoA^{WT} or Mut^{P2} differed only by a factor of 3, strongly suggesting that the pre-equilibrium accounts for it. Tertiary complex formation was characterized by time-resolved NMR^{21,22}. For Mut^{P2}, we found the folding rate constant (k_{F}) = $(10.0 \pm 0.9) \times 10^{-2} \text{ s}^{-1}$ compared to the wild-type k_{F} = $(4.0 \pm 0.1) \times 10^{-2} \text{ s}^{-1}$ (Fig. 2e). We further examined the functional differences for wild-type and Mut^{P2} in a transcription-translation-coupled luciferase assay (Fig. 3). Analysis of the normalized expression levels showed that Mut^{P2} was not able to control gene expression in an adenine-dependent manner but represented a constitutive on-state. By contrast, the wild type was able to increase expression levels in response to adenine addition. Comparison of the absolute expression levels of the wild type also revealed that the fourfold increase of gene expression over the entire temperature range correlated with a twofold lower overall expression level (Fig. 3).

For a two-state riboswitch, the temperature dependence of K_{d} makes regulation temperature-dependent for a given ratio of [ligand]:[riboswitch]. Our data for the *add* riboswitch from *V. vulnificus* strongly suggested that its regulation involves three distinct stable conformations. What additional function is conferred to the riboswitch by the

Errors were calculated from signal-to-noise (S/N) ratios of the NMR signals; identical results were obtained for two independent temperature and titration series at 800 MHz and 900 MHz. Mg^{2+} concentrations without adenine are 0, 2, 5 and 12 mM, and with adenine are 0, 5 and 12 mM (only for 10 °C and 30 °C). K_{pre} could be measured with NMR up to 30 °C only, as hydrogen exchange became too rapid at higher temperatures. **d**, Binding of adenine to wild-type monitored by ITC. The curves for binding of wild-type full-length RNA to adenine were recorded at 10 °C and 30 °C with 2 mM and 12 mM Mg^{2+} . Calculated $\log_{10} K_{\text{d}}$ is given for each binding curve. **e**, Conformational exchange in the apo state monitored by ^{15}N ZZ-chemical exchange spectroscopy for the wild-type RNA at 25 °C without Mg^{2+} . Black curves describe the decay of the diagonal peaks G44^{apoA} and G44^{apoB} with the mixing time. Blue and red curves describe the build-up and decay of the cross peaks between G44^{apoA} and G44^{apoB}. Cross peak intensities are scaled by a factor of 5 in the figure. Errors were calculated from S/N ratios of the NMR signals used in the analysis, exchange experiments were measured as duplicates and with three different relaxation delays ranging from 2 to 5 s, yielding identical results.

For a two-state riboswitch, the temperature dependence of K_{d} makes regulation temperature-dependent for a given ratio of [ligand]:[riboswitch]. Our data for the *add* riboswitch from *V. vulnificus* strongly suggested that its regulation involves three distinct stable conformations. What additional function is conferred to the riboswitch by the

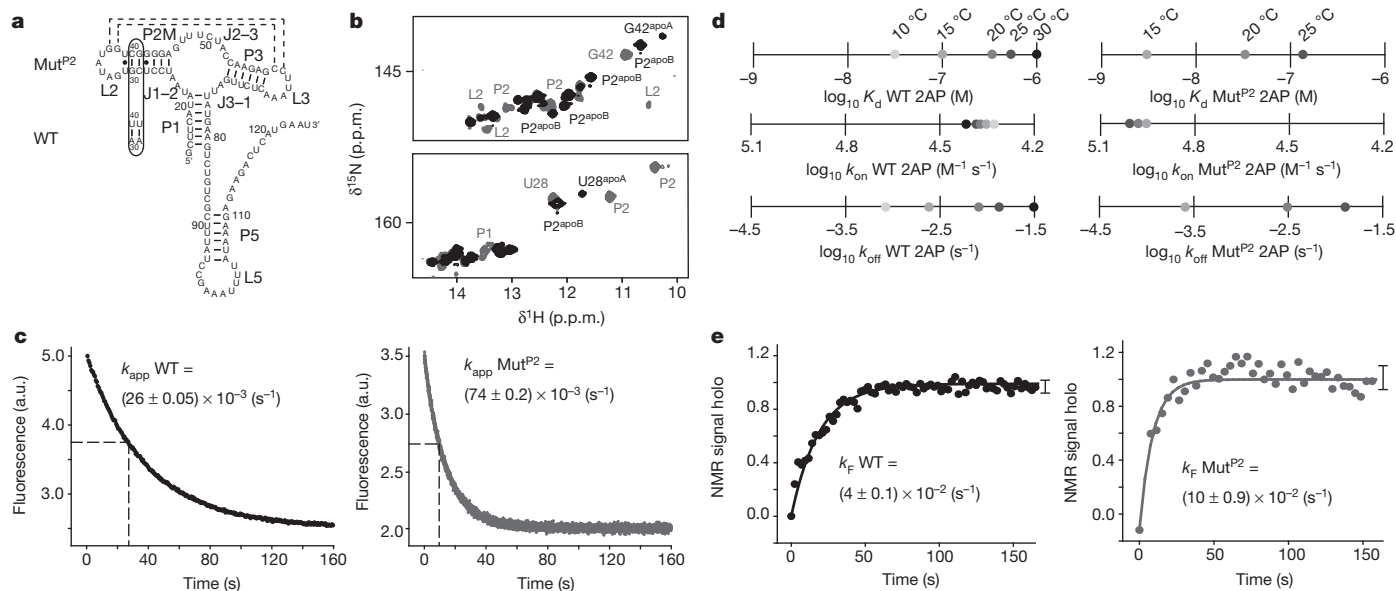


Figure 2 | Comparison of ligand binding kinetics of WT and Mut^{P2} adenine-sensing riboswitch constructs. **a**, Secondary structure of Mut^{P2} RNA in its apo conformation without adenine. Mutation site (A29C-U41G and A30G-U40C) is highlighted in comparison to the wild-type (WT) RNA construct. All detected secondary structure elements are shown. **b**, ¹H-¹⁵N HSQC spectra indicate the structural differences in the apo states for wild type (black) and Mut^{P2} (grey) in the absence of adenine and Mg²⁺ (20 °C, 900 MHz). Signals arising from conformation apoB are only visible for the wild type. The differences in the overlaid spectra are caused by the mutations in P2. New signals arising from G41 and G30 are visible, whereas those from U40 and U41 are absent. Signals next to the mutation site (U31/U39 and U28/G42) are shifted. Notably different to the wild-type spectra, reporter signals are observable for the loop-loop interactions, indicating the formation of a compact and preformed conformation. **c**, Fluorescence quenching of 2-aminopurine reports on the ligand binding to the adenine-sensing riboswitch. Normalized stopped-flow traces show ligand binding for wild-type (black) and Mut^{P2} (grey), 800 nM RNA, 2 mM Mg²⁺, 20 °C. Apparent rate constants k_{app} revealed faster binding kinetics for Mut^{P2} than for wild type. Shown are original curves, for the calculation of binding parameters, as depicted in **d** and tabulated in Supplementary Tables 3 and 4. These experiments were independently repeated at least three times, the errors given for the apparent rates are standard errors representing the deviation of the fit

from the experimental values. **d**, Stopped-flow binding kinetics, plots of $\log K_d$, $\log k_{on}$ and $\log k_{off}$ values of wild-type and Mut^{P2} RNA for five and three different temperatures, respectively. The K_d shows a more pronounced temperature dependence for Mut^{P2}. Although k_{on} is only marginally temperature-dependent for both wild type and Mut^{P2}, k_{on} of Mut^{P2} is one order of magnitude larger than k_{on} of the wild type. By contrast, the k_{off} values have a higher temperature dependence but are similar for wild type and Mut^{P2}. The size of error bars, calculated from the mean deviation of triplicate measurements, is smaller than the size of the marks. **e**, Monitoring formation of the holo conformation after ligand binding by real-time NMR. Sum of normalized integrals of imino proton signals from the holo conformation (U31, U39, U47, U49) as a function of time after rapid mixing of the RNA (0.3 mM, 2 mM Mg²⁺) with adenine (0.6 mM, 2 mM Mg²⁺) at 20 °C. The displayed error bars were calculated from S/N ratios of the peaks in the NMR spectra; the wild-type and Mut^{P2} experiments were performed as quadruplicates and duplicates, respectively. Data were fitted mono-exponentially to calculate an apparent rate constant for the formation of the holo conformation k_F . The errors given for the folding rates are standard errors representing the deviation of the fit from the experimental values. Similar to the binding kinetics, complex tertiary structure formation is faster for Mut^{P2} than for the wild type.

pre-equilibrium, with the potential consequence that the cell retains a constitutively inactive fraction of the *add* messenger RNA²³?

To address this question, we simulated the temperature-dependent change in conformational switching at cellular RNA concentrations of

1 nM²⁴, according to equation (2) (see Methods) (Fig. 4c, d). We introduced the term switching efficiency (equation (3), Methods), given by the difference in the populations of the holo conformation at two ligand concentrations. Although the cellular concentration of adenine in *Escherichia coli* has been reported to be in the range of 1.5 μ M²⁵, the differences in ligand concentration at which regulation occurs are unknown. We therefore considered a change in ligand concentration

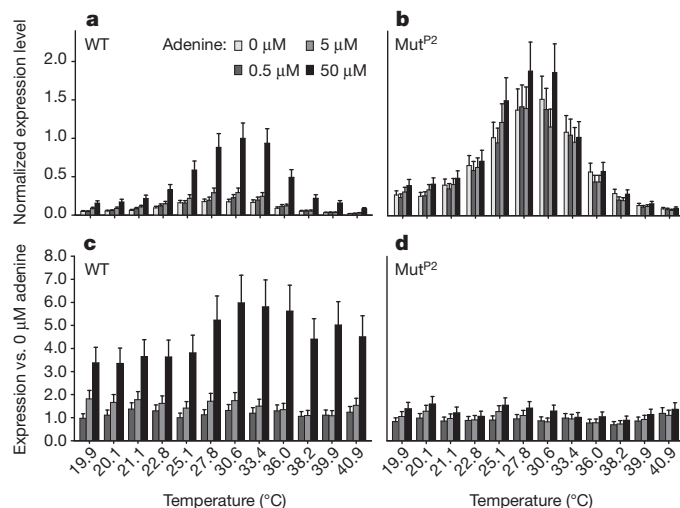


Figure 3 | Adenine-dependent expression regulation is only detected for the wild-type riboswitch. Transcription-translation-coupled *in vitro* luciferase assays at adenine concentrations from 0 to 50 μ M between 20 °C and 41 °C. **a**, **b**, Expression levels of wild type (**a**) and Mut^{P2} (**b**), normalized against the highest wild-type raw value (30.6 °C, 50 μ M). The overall expression level of Mut^{P2} showed a twofold increase compared to wild type. Error bars represent the standard deviation of three independent measurements. **c**, **d**, Expression levels of wild type (**c**) and Mut^{P2} (**d**), normalized against the respective values at 0 μ M adenine. At 50 μ M adenine, expression levels of wild type were increased over the entire temperature range by up to sixfold, whereas Mut^{P2} showed no adenine-dependent increase beyond experimental error, representing a constitutive on-state. Adenine independence was also observed in the expression of a luciferase control plasmid lacking either upstream riboswitch (Supplementary Fig. 10). Error bars are calculated by error propagation from the standard deviation of three independent measurements shown in **a** and **b**.

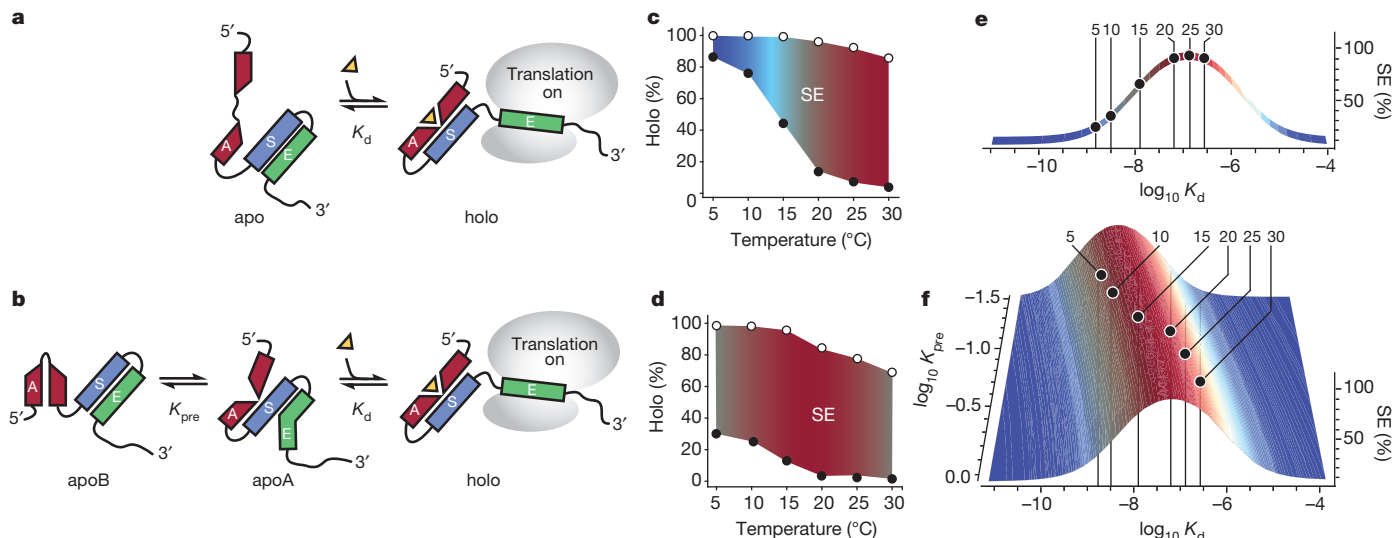


Figure 4 | Ligand-dependent conformational equilibria and simulation of the switching efficiency. **a, b**, Schematic picture of the two-state model (**a**) and the three-state model proposed here (**b**) for riboswitch regulation. Structural changes of aptamer (A), switching sequence (S) and effector sequence (E) are shown. **c**, Simulation of the temperature-dependent changes of the holo population for two adenine concentrations (0.01 μM (filled circles) and 1.5 μM (open circles)) at cellular mRNA concentrations of 1 nM for a two-state model. The switching efficiency (SE) is the area between the two holo populations for each temperature. **d**, Simulation of the temperature-dependent changes of the holo population for two adenine concentrations (0.01 μM (filled circles) and 1.5 μM (open circles)) at cellular mRNA concentrations of 1 nM for a three-state model. The switching efficiency is the area between the two holo

populations for each temperature. **e**, Simulation of the switching efficiency for riboswitches regulating gene expression with a two-state mechanism. The dependence of the switching efficiency on temperature without pre-equilibrium is shown. In the simulations, the adenine concentration changes from 0.01 to 1.5 μM and the RNA concentration is 1 nM. **f**, Simulation of the switching efficiency for riboswitches regulating gene expression with a three-state mechanism. Simulations were performed assuming the experimentally determined K_d and K_{pre} and changes in adenine concentration from 0.01 to 1.5 μM at a RNA concentration of 1 nM. The switching efficiency is constant over the given temperature range. All simulations are based on data measured at 2 mM Mg^{2+} .

from 0.01 to 1.5 μM for the two-state and three-state mechanism. The temperature-dependent changes of K_d significantly affected the switching efficiency for a two-state riboswitch: we found an increase in switching efficiency from 14% to 85% for a change in temperature from 5 to 30 $^{\circ}\text{C}$ (Fig. 4c, e). Owing to the pre-equilibrium in the wild type, the riboswitch was able to maintain a switching efficiency between 67% and 83% in the same temperature range (Fig. 4d, f). At low temperatures, the affinity of the riboswitch to adenine was high ($\log K_d = -8.5$) and at the same time, the pre-equilibrium was shifted towards the inactive apoB conformation. At higher temperatures, a potentially lower switching efficiency due to lower affinity ($\log K_d = -6.6$) was counteracted by a shift in the pre-equilibrium towards a higher population of apoA.

In conclusion, we found that three conformations, linked by two equilibria, were functionally relevant for the translational adenine-sensing riboswitch. This three-state behaviour (Fig. 4b) maintained constant switching efficiency for a larger variation in temperature and a tighter dependence on ligand concentration than a two-state riboswitch (Fig. 4a) could achieve. The temperature-responsive mechanism was clearly distinct from typical zipper-like thermometers^{26,27}, and might have an important role in *Vibrio* biology when the pathogen is transferred to humans from its marine habitat. In fact, genes in *V. vulnificus* encoded on chromosome II have been suggested to adapt to environmental conditions²⁸. The findings, however, have more general significance. By characterizing the sequence requirements for a three-state switch, we provided important information for synthetic biology applications of riboswitches. Furthermore, the three-state mechanism, also evidenced in conformational selection processes during ligand binding for a second riboswitch²⁹, is applicable in biomolecular assemblies in general. The ability to sense variations in temperature and ligand concentration simultaneously leads us to propose the term ‘riboswitch-thermostat’ for this new function in regulatory RNA elements.

METHODS SUMMARY

The unlabelled and ^{15}N -labelled RNA constructs were synthesized using T7 polymerase. All NMR experiments were performed in NMR buffer (25 mM potassium phosphate, 50 mM potassium chloride, pH 6.2) containing 10% D_2O on Bruker spectrometers and analysed with Topspin. To determine k_{AB} and k_{BA} , two-dimensional ^1H - ^{15}N heteronuclear exchange experiments were recorded on the full-length RNA without adenine. To record kinetics of the tertiary complex formation, NMR real-time kinetic experiments of the wild-type and Mut^{P2} RNA were conducted using a rapid sample mixing device. Binding affinities of adenine to wild-type full-length RNA were determined by ITC on a microcalorimeter (Microcal) in NMR buffer at different Mg^{2+} concentrations and temperatures. Stopped-flow kinetics were measured on an Applied Photophysics π^* -180 fluorescence spectrometer. The experiments were measured under pseudo-first order conditions in NMR buffer containing 2 mM Mg^{2+} . The fluorescent analogue of adenine, 2-aminopurine, was used. The inverted apparent rate constant $1/\tau$ was plotted against the RNA concentration to determine k_{on} , k_{off} and K_d values using equation (1)

$$\frac{1}{\tau} = k_{on}[\text{RNA}_0] + k_{off} \quad (1)$$

Circular dichroism melting curves were recorded on a J-810 CD spectrometer (Jasco) in NMR buffer without Mg^{2+} . Luciferase *in vitro* expression assays were performed on a gradient thermocycler at different temperatures and at four different adenine concentrations using DNA templates with wild-type, Mut^{P2} or no riboswitch cloned upstream of the firefly luciferase start codon. Expression levels were quantified on a microplate luminometer.

Theoretical simulations were performed using scripts written in the program Mathematica. The population of the holo conformation was calculated using equation (2)

$$P[\text{holo}] = \frac{100}{2[\text{RNA}_0]K_{pre}} \left(\frac{K_d + K_d K_{pre} + K_{pre}[L_0] + K_{pre}[\text{RNA}_0] - \sqrt{-4K_{pre}^2[L_0][\text{RNA}_0] + (K_d + K_d K_{pre} + K_{pre}[L_0] + K_{pre}[\text{RNA}_0])^2}}{(K_d + K_d K_{pre} + K_{pre}[L_0] + K_{pre}[\text{RNA}_0])^2} \right) \quad (2)$$

Full Methods and any associated references are available in the online version of the paper.

Received 25 April; accepted 12 June 2013.

Published online 10 July 2013.

- Mironov, A. S. *et al.* Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **111**, 747–756 (2002).
- Winkler, W., Nahvi, A. & Breaker, R. R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952–956 (2002).
- Nahvi, A. *et al.* Genetic control by a metabolite binding mRNA. *Chem. Biol.* **9**, 1043–1049 (2002).
- Garst, A. D., Edwards, A. L. & Batey, R. T. Riboswitches: structures and mechanisms. *Cold Spring Harb. Perspect. Biol.* **3**, a003533 (2011).
- Winkler, W. C. & Breaker, R. R. Genetic control by metabolite-binding riboswitches. *ChemBiochem* **4**, 1024–1032 (2003).
- Mandal, M. & Breaker, R. R. Adenine riboswitches and gene activation by disruption of a transcription terminator. *Nature Struct. Mol. Biol.* **11**, 29–35 (2004).
- Wickiser, J. K., Cheah, M. T., Breaker, R. R. & Crothers, D. M. The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* **44**, 13404–13414 (2005).
- Lemay, J. F. & Lafontaine, D. A. Core requirements of the adenine riboswitch aptamer for ligand binding. *RNA* **13**, 339–350 (2007).
- Greenleaf, W. J., Frieda, K. L., Foster, D. A., Woodside, M. T. & Block, S. M. Direct observation of hierarchical folding in single riboswitch aptamers. *Science* **319**, 630–633 (2008).
- Lemay, J. F., Penedo, J. C., Tremblay, R., Lilley, D. M. & Lafontaine, D. A. Folding of the adenine riboswitch. *Chem. Biol.* **13**, 857–868 (2006).
- Noeske, J. *et al.* An intermolecular base triple as the basis of ligand specificity and affinity in the guanine- and adenine-sensing riboswitch RNAs. *Proc. Natl Acad. Sci. USA* **102**, 1372–1377 (2005).
- Lemay, J. F. *et al.* Comparative study between transcriptionally- and translationally-acting adenine riboswitches reveals key differences in riboswitch regulatory mechanisms. *PLoS Genet.* **7**, e1001278 (2011).
- Neupane, K., Yu, H., Foster, D. A., Wang, F. & Woodside, M. T. Single-molecule force spectroscopy of the add adenine riboswitch relates folding to regulatory mechanism. *Nucleic Acids Res.* **39**, 7677–7687 (2011).
- Rieder, R., Lang, K., Graber, D. & Micura, R. Ligand-induced folding of the adenosine deaminase A-riboswitch and implications on riboswitch translational control. *ChemBiochem* **8**, 896–902 (2007).
- Serganov, A. *et al.* Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11**, 1729–1741 (2004).
- Lee, M. K., Gal, M., Frydman, L. & Varani, G. Real-time multidimensional NMR follows RNA folding with second resolution. *Proc. Natl Acad. Sci. USA* **107**, 9192–9197 (2010).
- Wang, J. *et al.* A method for helical RNA global structure determination in solution using small-angle X-ray scattering and NMR measurements. *J. Mol. Biol.* **393**, 717–734 (2009).
- Dixon, N. *et al.* Reengineering orthogonally selective riboswitches. *Proc. Natl Acad. Sci. USA* **107**, 2830–2835 (2010).
- Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **4**, 727–734 (1994).
- Stoddard, C. D. *et al.* Nucleotides adjacent to the ligand-binding pocket are linked to activity tuning in the purine riboswitch. *J. Mol. Biol.* **425**, 1596–1611 (2013).
- Buck, J., Furtig, B., Noeske, J., Wöhnert, J. & Schwalbe, H. Time-resolved NMR methods resolving ligand-induced RNA folding at atomic resolution. *Proc. Natl Acad. Sci. USA* **104**, 15699–15704 (2007).
- Mok, K. H. *et al.* Rapid sample-mixing technique for transient NMR and photo-CIDNP spectroscopy: Applications to real-time protein folding. *J. Am. Chem. Soc.* **125**, 12484–12492 (2003).
- Li, Y. C., Li, Y. M., Zhang, H. & Chen, Y. MicroRNA-mediated positive feedback loop and optimized bistable switch in a cancer network involving miR-17–92. *PLoS ONE* **6**, e18954 (2011).
- Selinger, D. W. *et al.* RNA expression analysis using a 30 base pair resolution *Escherichia coli* genome array. *Nature Biotechnol.* **18**, 1262–1268 (2000).
- Bennett, B. D. *et al.* Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chem. Biol.* **5**, 593–599 (2009).
- Rinnenhal, J., Klinkert, B., Narberhaus, F. & Schwalbe, H. Direct observation of the temperature-induced melting process of the *Salmonella* fourU RNA thermometer at base-pair resolution. *Nucleic Acids Res.* **38**, 3834–3847 (2010).
- Kortmann, J. & Narberhaus, F. Bacterial RNA thermometers: molecular zippers and switches. *Nature Rev. Microbiol.* **10**, 255–265 (2012).
- Xu, Q., Dziejman, M. & Mekalanos, J. J. Determination of the transcriptome of *Vibrio cholerae* during intrainestinal growth and midexponential phase *in vitro*. *Proc. Natl Acad. Sci. USA* **100**, 1286–1291 (2003).
- Wilson, R. C. *et al.* Tuning riboswitch regulation through conformational selection. *J. Mol. Biol.* **405**, 926–938 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank E. Stirnal, H. Keller and C. Richter for technical support. We thank F. Narberhaus, J. Wöhnert, J. Wachtveitl, J. Soppa, E. Schleiff, A. Heckel, M. Hengesbach and C. Griesinger for stimulating discussions. This work was funded by the German funding agency (DFG) in Collaborative Research Center 902.

Author Contributions A.R., S.N., K.S., F.B. and B.F. conducted experiments. All authors contributed to the analysis of the data and the writing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.S. (schwalbe@nmr.uni-frankfurt.de) or B.F. (furtig@nmr.uni-frankfurt.de).

METHODS

RNA preparation. The unlabelled and ^{15}N -labelled RNA constructs were synthesized by *in vitro* transcription using T7 polymerase and purified as described³⁰. All RNAs were folded by thermal denaturation of the RNA at high concentration (0.2–0.5 mM), diluted to 0.05 mM, and rapidly cooled on ice. Folding into a homogenous conformation was verified by native PAGE. Samples were exchanged into NMR buffer (25 mM potassium phosphate, 50 mM potassium chloride, pH 6.2). A 5'-hammerhead ribozyme³¹ was used for the transcription of wild-type and Mut^{P2} RNAs.

Synthesis of labelled adenine. The ^{13}C , ^{15}N -labelled adenine was synthesized and analysed as described previously¹¹. The concentration of the labelled adenine was estimated using ultraviolet absorbance at 261 nm, using $\epsilon_{261\text{nm}} = 13,400 \text{ mol}^{-1} \text{ cm}^{-1}$ (ref. 32).

NMR spectroscopy. NMR experiments were conducted on Bruker spectrometers (AV600-AV950). Processing and analysis of the data were performed using the software programs topspin 1.3-3.1 (Bruker BioSpin, Rheinstetten) and Sparky 3.114 (T. D. Goddard & D. G. Kneller, University of California, San Francisco). Samples contained 90% H_2O and 10% D_2O .

Assignment of secondary and tertiary structure elements within the apo and holo states of the RNA were based on the chemical shift assignments of imino proton resonances (Fig. 1). These resonances are reporters for the interaction between adjacent base pairs, and were assigned using NOESY (nuclear Overhauser enhancement spectroscopy) and HSQC experiments, and by comparing experiments of full-length RNAs with fragment structural reference modules (Supplementary Figs 2–4 and Supplementary Tables 1 and 2).

^{15}N ZZ-exchange. Two-dimensional ^1H - ^{15}N heteronuclear exchange experiments¹⁹ were recorded to determine the rate of interconversion between the apo states. For the ligand-free wild-type full-length RNA, nucleotides G43 and G44 yielded detectable imino proton signals in both folds (Supplementary Fig. 6a, b). Owing to spectral overlap, only the time dependence of peak heights of the diagonal and cross peaks of G44 (mixing times: 0–0.8 s) were quantified (Fig. 1d and Supplementary Fig. 6). Data points were fitted with Mathematica (Wolfram Research) to extract the chemical exchange rates k_{AB} and k_{BA} . The experiments were performed on an AV700 spectrometer with 0 and 2 mM Mg^{2+} at 25 °C, and with 2 mM Mg^{2+} at 30 °C (Supplementary Fig. 6c).

Real-time NMR kinetics. NMR real-time kinetic experiments of the wild-type and Mut^{P2} RNA were conducted on an AV800 spectrometer at 20 °C with 2 mM Mg^{2+} . The reaction was initiated inside a shigemi-tube using a rapid sample mixing device²². Three-hundred microlitres of 0.3 mM RNA was mixed with 40 μl of 5 mM adenine solution to prepare an $[\text{RNA}]:[\text{ligand}]$ ratio of 1:2. For the wild-type construct, a selectively ^{15}N -uridine-labelled RNA in combination with ^{15}N -filtered/edited experiments^{21,33} was used. For the Mut^{P2} construct, a fully labelled RNA was used. The experiments were recorded as a pseudo-2D dataset³⁴. The overall apparent wild-type and Mut^{P2} k_{f} rate constants for folding of the complexes were determined by calculating the sum of all normalized imino proton signal integrals that showed time-dependent changes and were well resolved. The obtained kinetic traces were fitted with a mono-exponential function.

ITC RNA-binding assays. The binding affinities between adenine and full-length wild-type RNA were measured using a microcalorimeter (Microcal). The RNA (40 μM) and adenine (100 μM) were dissolved in NMR buffer containing 2 mM Mg^{2+} . Measurements were performed at 10 °C and 30 °C. K_{d} (equilibrium binding constant) values were calculated by fitting the raw ITC data with nonlinear least-squares analysis (Fig. 1d, e).

Stopped-flow kinetics. Stopped-flow kinetics were measured on an Applied Photophysics π^* -180 fluorescence spectrometer. A fluorescent analogue of adenine, 2-aminopurine, binds to the adenine-sensing riboswitch aptamer, similar to the cognate ligand adenine^{6,10,15}. RNA and 2-aminopurine were dissolved in NMR buffer containing 2 mM MgCl_2 . The 2-aminopurine concentration was calculated using ultraviolet absorbance at 305 nm³². We performed the stopped-flow measurements and analysis as previously described⁷. The RNA concentration was varied from 100 nM to 1 mM in excess over 2-aminopurine (50 nM). Rate constants

were determined at 10, 15, 20, 25 and 30 °C for the wild-type RNA, and at 15, 20 and 25 °C for Mut^{P2} RNA. The fluorescence decay was fitted to a mono-exponential decay with three parameters (Supplementary Fig. 9). The resulting apparent rate constants used for analysis are averaged values of at least three independent measurements (Supplementary Table 3). Association (k_{on}) and dissociation (k_{off}) rate constants were determined by plotting the inverted apparent rate constant $1/\tau$ versus the RNA concentration ($[\text{RNA}_0]$) using equation (1) (Supplementary Table 4):

$$\frac{1}{\tau} = k_{\text{on}}[\text{RNA}_0] + k_{\text{off}} \quad (1)$$

Circular dichroism melting. Circular dichroism melting curves were recorded on a J-810 CD spectrometer (Jasco) using a 1-mm quartz cuvette, with 250 μl volume at a wavelength of 264 nm. The RNA concentration was 5 μM in NMR buffer. Spectra were recorded with 10 μM adenine and without Mg^{2+} . Circular dichroism melting curves were analysed according to the literature^{35,36}.

Luciferase *in vitro* expression assays. DNA constructs of wild-type and Mut^{P2} were designed as follows. The triplet A120-U121-G122 downstream of the ribosome-binding site of the adenine riboswitch constituted the start codon of a firefly luciferase reporter gene. The constructs were cloned into the plasmid pET23(+), which carries no intrinsic ribosome-binding site. For negative control experiments, a T7 luciferase control plasmid (Promega) was used (Supplementary Fig. 11).

For the assay, 15 μg of template DNA were added to a total volume of 200 μl of a transcription-translation-coupled *E. coli*-based *in vitro* expression system (RTS 100 *E. coli* HY, 5 PRIME GmbH), supplemented with 250 U of RNase inhibitor (RiboLock, Thermo Scientific). One microlitre of adenine stock solutions of 25 mM, 2.5 mM and 250 μM , respectively, were added to 50- μl aliquots of the reaction mix, which were then further divided into 4- μl aliquots and incubated on a gradient thermocycler (Eppendorf AG) for 1 h with a temperature gradient of 20–41 °C.

Samples (2.5 μl) of each aliquot were then added to 50 μl of SteadyGlo luciferase substrate solution (Promega) on a microtitre plate. After 5 min incubation at room temperature, bioluminescence was quantified using a Veritas Microplate Luminometer (Turner BioSystems). All experiments were performed in triplicates, with a total Mg^{2+} concentration of 12 mM.

Simulations. Simulation of the population of the holo RNA conformation ($P[\text{holo}]$) and switching efficiency (SE) were performed using scripts written in the program Mathematica using equations (2) and (3):

$$P[\text{holo}] = \frac{100}{2[\text{RNA}_0]K_{\text{pre}}} \left(\frac{K_{\text{d}} + K_{\text{d}}K_{\text{pre}} + K_{\text{pre}}[\text{L}_0] + K_{\text{pre}}[\text{RNA}_0] - \sqrt{-4K_{\text{pre}}^2[\text{L}_0][\text{RNA}_0] + (K_{\text{d}} + K_{\text{d}}K_{\text{pre}} + K_{\text{pre}}[\text{L}_0] + K_{\text{pre}}[\text{RNA}_0])^2}}{K_{\text{d}} + K_{\text{d}}K_{\text{pre}} + K_{\text{pre}}[\text{L}_0] + K_{\text{pre}}[\text{RNA}_0]} \right) \quad (2)$$

$$\text{SE}(L_{0,1}, L_{0,2}, \text{RNA}_0) = P[\text{holo}](L_{0,1}, \text{RNA}_0) - P[\text{holo}](L_{0,2}, \text{RNA}_0) \quad (3)$$

Derivation of the equations is detailed in Supplementary Information.

30. Stoldt, M., Wohnert, J., Ohlenschläger, O., Goriach, M. & Brown, L. R. The NMR structure of the 5S rRNA E-domain-protein L25 complex shows preformed and induced recognition. *EMBO J.* **18**, 6508–6521 (1999).
31. Birikh, K. R., Heaton, P. A. & Eckstein, F. The structure, function and application of the hammerhead ribozyme. *Eur. J. Biochem.* **245**, 1–16 (1997).
32. Fasman, G. D. *Handbook of Biochemistry and Molecular Biology, Nucleic Acids* 3rd edn, Vol. 1, 65–215 (CRC Press, 1975).
33. Piotto, M., Saudek, V. & Sklenar, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomolec. NMR* **2**, 661–665 (1992).
34. Wacker, A., Buck, J., Richter, C., Schwalbe, H. & Wohnert, J. Mechanisms for differentiation between cognate and near-cognate ligands by purine riboswitches. *RNA Biol.* **9**, 672–680 (2012).
35. Marky, L. A. & Breslauer, K. J. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**, 1601–1620 (1987).
36. Mergny, J. L. & Lacroix, L. Analysis of thermal melting curves. *Oligonucleotides* **13**, 515–537 (2003).

BACTERIOLOGY

Toxins in tandem

Analyses of the toxin produced by *Salmonella* Typhi bacteria reveal an unusual assembly of toxin subunits, and show that most symptoms of typhoid fever can be linked to one subunit's DNA-damaging activity. SEE LETTER P.350

C. EREC STEBBINS

During the complex process of infection, nothing is more fascinating than the interplay between host immunity and pathogen virulence — particularly in asymptomatic carriers, who seem to be in perfect health while normally life-threatening organisms replicate within them. The most infamous such carrier was 'Typhoid Mary', a cook in the United States who infected more than 50 people with typhoid fever by passing on the causative bacterium *Salmonella* Typhi (Fig. 1). Another perplexing feature of these bacteria is the existence of the closely related *Salmonella* Typhimurium, which does not cause life-threatening infections despite having apparently similar virulence properties to *S. Typhi*. A paper by Song *et al.*¹, on page 350 of this issue, places several key pieces in this puzzle.*

The story centres on one of the few obvious differences between *S. Typhi* and *S. Typhimurium*: the typhoid toxin, which was discovered almost 10 years ago and is specific to *S. Typhi*^{2–4}. Typhoid toxin has several features that set it apart from other bacterial toxins. It is known to be assembled from three subunits: CdtB and PltA (proteins with similarities to two disparate classes of bacterial toxins, the cytolethal distending genotoxins (CDT) and the ADP-ribosyl transferases of the pertussis toxin family, respectively), and PltB. But how this dual-activity toxin contributes to the pathogenesis of typhoid fever was, until now, unclear.

Song and colleagues show that, in mice, typhoid toxin alone is sufficient to induce most of the symptoms associated with *S. Typhi* infections, and that these effects are mediated by the DNA-damaging CdtB subunit of the toxin. The authors also show that the toxin is targeted to host cells by interactions with carbohydrate moieties on specific surface glycoproteins. This is a remarkable finding, because previous work had failed to provide a *raison d'être* for CDT toxins in the pathogenesis of other bacteria that carry such proteins. The discovery that the symptoms of typhoid

*This article and the paper under discussion¹ were published online on 10 July 2013.



Figure 1 | Typhoid Mary. A classic image of Mary Mallon, an asymptomatic carrier of *Salmonella* Typhi, from public-health posters of the early 1900s. Mary was a cook in New York state who lived a symptom-free life but infected more than 50 people with typhoid fever before health-department detectives traced the epidemics and moved her to involuntary quarantine.

fever depend on CdtB of *S. Typhi* is the first time that the virulence of an organism has been directly linked to the ability of bacteria to induce cellular DNA damage in a host.

The second amazing insight from this study involves the architecture of the typhoid toxin, as revealed by the X-ray crystal structure described by the authors. The structure shows typhoid toxin to be a classic 'AB' type in its general assembly, possessing an active (A) subunit that achieves a biochemical result when a binding (B) subunit interacts with specific cellular receptors. The B subunit of typhoid toxin (PltB) is a homologue of one of the heteropentameric subunits of the pertussis toxin B subunits and, as expected, it pentamerizes with other B subunits. The PltA subunit nestles tightly in the cup that forms at one side of this pentameric disk (see Fig. 3 of the paper¹).

However, in a striking departure from other known toxin structures, the CdtB subunit of typhoid toxin is not tightly integrated into the main structure, but is 'glued' onto it by a single disulphide bond to the PltA subunit. The authors find that this disulphide bond is crucial for toxin assembly, and that it creates an A₂B₅ organization that has not been seen in any other bacterial toxin. This structure provides two disparate enzymatic activities, only one of which (the DNA-damaging activity) is required to recapitulate most typhoid symptoms.

This unusual tethering of CdtB to PltA could represent a 'missing link' in toxin evolution, in which a factor that confers an advantage to an existing toxin is crudely conjugated in order to associate two cellular activities. Perhaps in several million years, a better-integrated multicomponent toxin will have replaced this one, but for now it is fascinating to obtain this snapshot of evolution in action.

Much remains to be explained. For example, the role of the PltA subunit is not yet accounted for, and although typhoid toxin alone can reproduce many symptoms of typhoid disease, it does not cause the associated fever. However, this work, with its sweeping combination of atomic-resolution reductionism and organismal infection studies, has identified some of the key elements that underlie the unique properties of this pathogen. Our enhanced understanding of the role of the CdtB subunit in *S. Typhi* virulence may open doors to new pharmacological interventions and improved vaccines for an infection for which there are no robust immunization approaches. Typhoid still takes the lives of several hundred thousand people a year^{5–7}, and the rise of antibiotic-resistant strains of *S. Typhi* could take us back to a situation in which individuals like Typhoid Mary pose a public-health threat. Research

to improve our understanding of the disease and to develop treatments is an important undertaking. ■

C. Erec Stebbins is in the Laboratory of Structural Microbiology, The Rockefeller University, New York, New York 10065, USA. e-mail: stebbins@rockefeller.edu

1. Song, J., Gao, X. & Galán, J. E. *Nature* **499**, 350–354 (2013).
2. Haghjoo, E. & Galán, J. E. *Proc. Natl Acad. Sci. USA* **101**, 4614–4619 (2004).
3. Spano, S., Ugalde, J. E. & Galán, J. E. *Cell Host Microbe* **3**, 30–38 (2008).
4. Spano, S. & Galán, J. E. *Curr. Opin. Microbiol.* **11**, 15–20 (2008).
5. Parry, C., Hien, T. T., Dougan, G., White, N. & Farrar, J. *N. Engl. J. Med.* **347**, 1770–1782 (2002).
6. Raffatellu, M., Wilson, R., Winter, S. & Baumler, A. *J. Infect. Dev. Countries* **2**, 260–266 (2008).
7. Butler, T. *Clin. Microbiol. Infect.* **17**, 959–963 (2011).

9. Clark, D. A., Clark, D. B. & Oberbauer, S. F. *J. Geophys. Res.* **118**, 1–12 (2013).
10. Silva, L. C. R. & Anand, M. *Glob. Ecol. Biogeogr.* **22**, 83–92 (2013).
11. Donohue, R. J., McVicar, T. R. & Roderick, M. L. *Glob. Change Biol.* **15**, 1025–1039 (2009).
12. Baldocchi, D. *Aust. J. Bot.* **56**, 1–26 (2008).
13. Wong, S. C., Cowan, I. R. & Farquhar, G. D. *Plant Physiol.* **78**, 821–825 (1985).
14. Drake, B. G., González-Meler, M. A. & Long, S. P. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **28**, 609–639 (1996).
15. Ainsworth, E. A. & Rogers, A. *Plant Cell Environ.* **30**, 258–270 (2007).
16. Barton, C. V. M. *et al. Glob. Change Biol.* **18**, 585–595 (2012).
17. De Kauwe, M. G. *et al. Glob. Change Biol.* **19**, 1759–1779 (2013).

RNA BIOPHYSICS

A three-state balancing act

How do pathogens survive temperature variations? At a molecular level, one bacterial species seems to regulate gene expression in response to temperature through structural equilibria in corresponding RNA sequences. [SEE LETTER P.355](#)

RONALD MICURA

In bacteria, many messenger RNA molecules carry a regulatory segment called a riboswitch. Specific binding of small ligand molecules, such as adenine, to this segment determine whether the riboswitch mRNA will be translated into a protein^{1–4}. On page 355 of this issue, Reining *et al.*⁵ show that regulation of gene expression through such ‘riboswitching’ is coupled to temperature sensing. The authors investigate the adenine-sensitive riboswitch from the pathogenic bacterium *Vibrio vulnificus*⁶, and demonstrate that efficient RNA regulation at different temperatures — those of the bacterium’s marine habitat and its human host — requires a change in the riboswitch’s structural behaviour from a two-state pattern to a three-state one.*

A typical riboswitch consists of two domains at the 5′ end of the mRNA: a ligand-binding aptamer and an adjoining expression platform, which can have one of two mutually exclusive structures depending on whether the aptamer is in the ligand-bound or ligand-unbound state. The structural change in the expression platform signals that gene expression should be turned on or off.

During gene transcription, a polymerase enzyme synthesizes first the aptamer and then the expression platform⁷. The sequential release of riboswitch domains from the polymerase is especially meaningful for transcription-controlling riboswitches, which act under kinetic control. To direct the folding of the expression platform, the aptamer domain of the growing RNA chain must bind rapidly to its ligand, which requires high ligand concentrations⁸; otherwise, the resulting full-length mRNA becomes trapped in a default fold that cannot respond to the ligand⁹. In

transcription-controlling riboswitches, the two mutually exclusive structures are generally referred to as terminator and antiterminator folds, causing cessation of polymerase activity and continuation of mRNA synthesis, respectively.

Bacterial transcription is tightly coupled to translation. In translation-controlling riboswitches, the molecular mechanism relies on either sequestration or liberation of the Shine–Dalgarno sequence — the mRNA site that binds to cellular organelles known as ribosomes to initiate translation. In contrast to transcription-controlling riboswitches, most translation-controlling riboswitches act under thermodynamic control⁹. Consequently, ligand-dependent control of translation is maintained even for a full-length mRNA. For these riboswitches, therefore, two states (ligand-bound and -unbound) seem sufficient to turn translation on and off.

Reining and colleagues have found that, for robust functioning, their translation-controlling adenine-sensing riboswitch must occur in three structural states. The authors investigated the full-length (more than 100-nucleotide) riboswitch domain at single-nucleotide resolution. They also determined a complete set of thermodynamic and kinetic parameters for folding and ligand-binding of this RNA, under conditions that involved varying the concentrations of RNA, magnesium ions (a factor mediating structure formation) and adenine, as well as, importantly, temperature.

The researchers find that a ligand-free (apo) form of the riboswitch exists in a pre-equilibrium of two structurally distinct aptamer folds (Fig. 1). One of the structures (apoA) can bind to the ligand, exhibiting a structure that resembles a third structural state — the adenine-bound holo form. The other structure (apoB) adopts a different fold and cannot interact with the ligand. But why does



50 Years Ago

A technique has been used with limited success to obtain simultaneous measurement of cosmic radiation at two different altitudes. The method consists of suspending two packets of nuclear emulsion plates from the one balloon, while maintaining a constant vertical separation of 10,000–27,000 ft. between the packets ... in this way the two packets do not separate in latitude and longitude—a factor which enters if two separate independent balloons are flown.

This has been achieved by carrying aloft a cone of nylon string which was allowed to unwind at a predetermined height, leaving one packet of plates suspended at the balloon while the other fell at the end of the string ... As the alarm rings, the alarm winder releases the key ... holding the lower packet and target, which then fall away freely, unwinding the string in the process. Twenty thousand feet of string unwinds in approximately 10 min. **From *Nature* 20 July 1963**

100 Years Ago

The Potato: A Compilation of Information from Every Available Source. By E. H. Grubb and W. S. Guilford.

There are men who, having attained to wealth and fame by the agency of some humble instrument, basely repudiate and kick over the ladder by which they have risen. Not so the authors of the first book on our list. The potato has “made” them, and in return they proceed to “make” the potato ... the authors are so evidently enthusiastic, and discourse so eloquently on the merits of their subject, that we are carried along with them, and forget that, after all, they are only talking about potatoes, and not about alpine plants or roses.

From *Nature* 17 July 1913

*This article and the paper under discussion⁵ were published online on 10 July 2013.

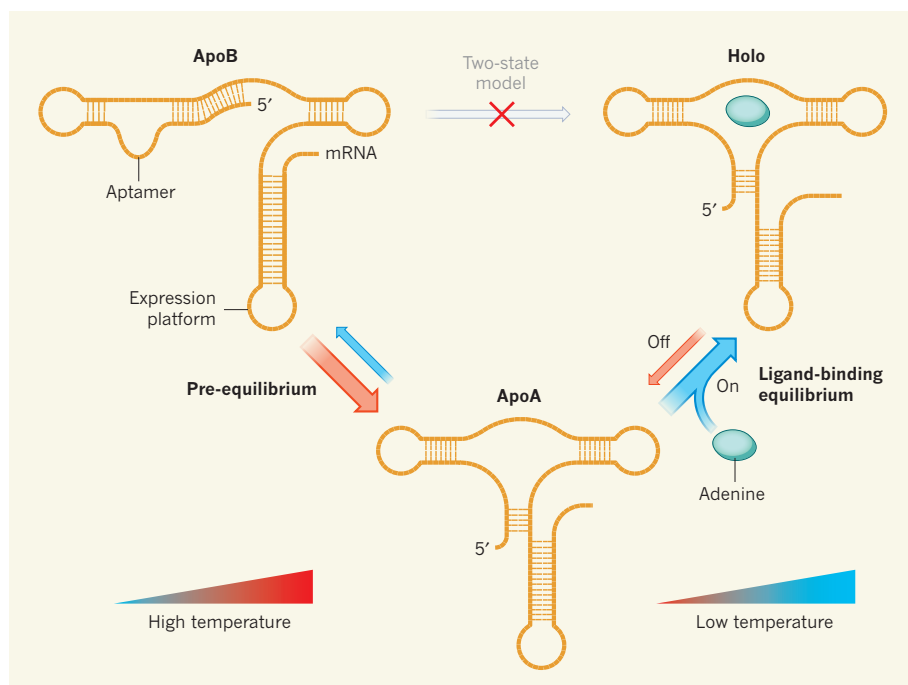


Figure 1 | Three states of a riboswitch. Riboswitches consist of an aptamer and an expression platform. Reining *et al.*⁵ find that, in contrast to the accepted two-state model, a three-state model exists for the structural behaviour of an adenine-sensing bacterial riboswitch. They show that the ligand-free apo form of this riboswitch can exist in two pre-equilibrium states: an apoA state that can bind a ligand (adenine), and an apoB state that cannot. On binding adenine, the apoA structure adopts a third (holo) state. The structural pre-equilibrium between apoA and apoB states counterbalances the temperature-dependent equilibrium between the apoA and holo states, guaranteeing effective regulation, over a broad temperature range, of the gene corresponding to the RNA sequence that carries the riboswitch.

this bistable sequence element exist?

Reining *et al.* provide a clear answer: this pre-equilibrium is markedly temperature dependent and counterbalances the temperature-dependent changes in ligand affinity. At low temperatures, when adenine has a high affinity for the RNA, only a small population of the adenine-binding apoA form is available, but it is sufficient to achieve high riboswitching efficiency (Fig. 1). At higher temperatures, this population is significantly increased and compensates for the lower RNA affinity of the ligand. Thus, whereas in a pure two-state model switching efficiency would vary drastically with environmental changes, in this three-state model the switching efficiency remains robust.

The authors further perform a coupled transcription–translation assay to underline the significance of their biophysical findings. They show that a mutant riboswitch that cannot adopt the apoB structure and exists only in the binding-competent apoA conformation lacks sensitivity to changes in adenine concentration and so cannot control gene expression in response to varying ligand doses. This mutant is also insensitive to changes in temperature and is always in an ‘on’ state.

Nature probably contains further examples of three-state structural behaviour for RNA that balances RNA-regulated gene expression. Bistable elements are widespread in RNA

sequences and have been thoroughly characterized in at least one other riboswitch system — a preQ₁ class I riboswitch from the human pathogen *Fusobacterium nucleatum*¹⁰. Moreover, temperature-sensitive RNAs that control gene expression, ‘RNA thermometers’¹¹, are prevalent. Coupling ligand and temperature sensitivities seems logical, although the three-state mechanism discovered by Reining *et al.*

PARTICLE PHYSICS

Let it B

A phenomenon known as CP asymmetry, which explains our very existence, has been observed in the decays of B_s^0 mesonic particles. The finding represents yet another triumph of the standard model of particle physics.

YOSEF NIR

The first observation of matter–antimatter asymmetry in the decays of particles known as B_s^0 mesons has been reported in *Physical Review Letters* by the LHCb collaboration¹ (Aaij *et al.*) at the Large Hadron Collider at CERN, near Geneva, Switzerland. This measurement reinforces the standard model of particle physics and, in particular, its

compensates for temperature fluctuations rather than using this parameter to directly trigger a gene response.

As the present paper shows, mechanistic insights into riboswitch function can provide a deeper understanding of the potential adaptive strategies used by bacteria in different environments. Such insights may also be of fundamental interest for exploring the possibility of targeting riboswitch RNAs with antibacterial drugs — for example, by interfering with the delicate balance of the RNA-structure equilibria. Furthermore, these findings will influence the design of RNA biosensors, which represent emerging tools for live-cell imaging¹². They will also affect the use of riboswitches in synthetic-biology applications for reprogramming cells to autonomously perform complex tasks, such as ‘seek-and-destroy’ herbicides¹³. ■

Ronald Micura is at the Institute of Organic Chemistry and the Center for Molecular Biosciences, Leopold Franzens University, 6020 Innsbruck, Austria.
e-mail: ronald.micura@uibk.ac.at

- Breaker, R. R. *Mol. Cell* **43**, 867–879 (2011).
- Serganov, A. & Patel, D. J. *Annu. Rev. Biophys. Biomol. Struct.* **41**, 343–370 (2012).
- Serganov, A. & Nudler, E. *Cell* **152**, 17–24 (2013).
- Liberman, J. A. & Wedekind, J. E. *Wiley Interdisc. Rev. RNA* **3**, 369–384 (2012).
- Reining, A. *et al. Nature* **499**, 355–359 (2013).
- Lemay, J.-F. *et al. PLoS Genet.* **7**, e1001278 (2011).
- Garst, A. D., Edwards, A. L. & Batey, R. T. *Cold Spring Harb. Perspect. Biol.* **3**, a003533 (2011).
- Al-Hashimi, H. M. & Walter, N. G. *Curr. Opin. Struct. Biol.* **18**, 321–329 (2008).
- Zhang, J., Lau, M. W. & Ferré-D’Amaré, A. R. *Biochemistry* **49**, 9123–9131 (2010).
- Rieder, U., Kreutz, C. & Micura, R. *Proc. Natl Acad. Sci. USA* **107**, 10804–10809 (2010).
- Kortmann, J. & Narberhaus, F. *Nature Rev. Microbiol.* **10**, 255–365 (2012).
- Kellenberger, C. A., Wilson, S. C., Sales-Lee, J. & Hammond, M. C. *J. Am. Chem. Soc.* **135**, 4906–4909 (2013).
- Sinha, J., Reyes, S. J. & Gallivan, J. P. *Nature Chem. Biol.* **6**, 464–470 (2010).

explanation of how the weak force, which governs radioactive decays, distinguishes matter from antimatter.

Matter particles and their antiparticles have opposite charge (C) and parity (P), the latter meaning that, under spatial-inversion transformation, the direction of a particle’s spin relative to its direction of motion is reversed. The difference in the laws of physics obeyed by matter and antimatter is experimentally

state of the doublet accommodates current in the 'forward' direction (the supercurrent usually observed in junctions), whereas the excited state corresponds to current in the 'reverse' direction. Importantly, the Andreev bound-state doublet is a system of spin-1/2 that is localized to the weak link and behaves as an internal degree of freedom associated with the Josephson effect. It is this internal feature that Bretheau *et al.* detected directly using the technique of photon-absorption spectroscopy.

The authors used several aspects of the Josephson effect to carry out photon-absorption spectroscopy of the Andreev doublet. First, to access the doublet from a single weak link, they used a mechanically controllable atomic break junction (a narrow bridge containing only a few atoms) formed on a flexible substrate⁵. By applying a mechanical force *in situ* to the cryogenically cooled break junction, the researchers could reduce the number of channels down to as low as a single weak link with a particular transmission probability. Second, to change the phase difference across the atomic junction and thereby tune the Andreev-doublet energies, Bretheau *et al.* fabricated the break junction as part of a superconducting loop containing a much larger Josephson tunnel junction, thereby forming a sensitive magnetometer known as a superconducting quantum interference device. In this configuration, a magnetic field applied to the loop sets the phase difference across the atomic junction and makes the bound-state energies tunable.

Third, to drive transitions between the ground and excited Andreev bound states, the authors used a separate tunnel junction as a spectrometer (Fig. 1). Applying a voltage to the spectrometer junction causes it to emit electromagnetic radiation at a frequency that is proportional to the applied voltage and is tunable over a wide range — up to around 80 gigahertz. Finally, owing to conservation of energy, a unidirectional (d.c.) current flows through the spectrometer junction that is proportional to the absorption rate of the radiation emitted by the atomic junction or the surrounding environment. Therefore, determining the spectrometer's d.c. current is a measurement of the absorption rate. The spectrometer Josephson junction thus serves both as a microwave generator and a detector in this experiment. With this set-up, Bretheau *et al.* carried out photon-absorption spectroscopy of the Andreev bound-state doublet, observing the expected trend in transition frequency as a function of the phase difference across the atomic junction.

It should be noted that in several earlier experimental works (for example, ref. 6), the presence of Andreev bound states was required to explain the observed phenomena, serving as indirect evidence for their existence. More recently, direct evidence was provided by applying a technique called tunnelling spectroscopy

to a carbon nanotube weak link⁷ and to a graphene quantum dot⁸ — complementary studies that observed alternative microscopic configurations of the bound states.

Bretheau and colleagues' experiment stands as the first photon-absorption spectroscopy of the spin-1/2 Andreev bound-state doublet. These doublets could be used as a quantum logic element or, in conjunction with interactions between the electron's spin and orbital degrees of freedom, to realize a Majorana state — an elusive spin-1/2 particle that is its own antiparticle. This experimental technique may also shed light on the perennial problem of quasi-particle poisoning in coherent superconducting devices. Indeed, by driving the bound-state transition, this work opens up experiments to a new level of Andreev physics. ■

BIOGEOCHEMISTRY

Carbon dioxide and water use in forests

Plants are expected to respond to rising levels of atmospheric carbon dioxide by using water more efficiently. Direct evidence of this has been obtained from forests, but the size of the effect will prompt debate. SEE LETTER P.324

BELINDA MEDLYN & MARTIN DE KAUWE

In a study published on page 324 of this issue, Keenan *et al.*¹ report that the efficiency with which forests use water has increased over the past 20 years, and conclude that this is a consequence of rises in the concentration of atmospheric carbon dioxide.* The findings call for a reassessment of models of the terrestrial carbon cycle.

The concentration of CO₂ in the atmosphere is rising at an unprecedented rate. In May this year, it reached 400 parts per million, 43% above the pre-industrial concentration of 280 p.p.m. (ref. 2). Much of this increase has occurred in recent decades, with the rate of increase over the past 20 years being 5% per decade². This drastic upsurge in atmospheric CO₂ should have stimulated plant productivity worldwide, because we know from experiments that rising CO₂ concentrations increase the rate of photosynthesis and reduce water use in plants³. Such effects are fundamental to our current understanding of the carbon cycle — for example, most terrestrial carbon-cycle models explain the current land sink for carbon by assuming that rising CO₂ levels have enhanced plant productivity⁴.

However, detecting the effects of rising

Simon Gustavsson is in the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139–4307, USA. **William D. Oliver** is in the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02420–9108, USA. e-mails: simongus@mit.edu; oliver@ll.mit.edu

1. Josephson, B. D. *Phys. Lett.* **1**, 251 (1962).
2. Bretheau, L. *et al. Nature* **499**, 312–315 (2013).
3. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* (Basic Books, 2011).
4. Hekking, F. W. J., Schon, G. & Averin, D. V. *Proc. NATO Adv. Res. Workshop Mesoscopic Superconductivity* (NATO-ASI, 1994).
5. van Ruitenbeek, J. M. *et al. Rev. Sci. Instrum.* **67**, 108–111 (1996).
6. Fueschle, M. *et al. Phys. Rev. Lett.* **102**, 127001 (2009).
7. Pillet, J.-D. *et al. Nature Phys.* **6**, 965–969 (2010).
8. Dirks, T. *et al. Nature Phys.* **7**, 386–390 (2011).

CO₂ concentrations on terrestrial vegetation outside controlled experiments has proven remarkably difficult, provoking numerous debates about whether such effects are really occurring^{5–10}. There are few high-quality, long-term records of plant productivity and water use that can be used to test for such effects. The main types of data come from plot surveys, tree-ring records, satellite images, aerial photographs and measurements of stream flow. Each of these is an indirect measurement and has a relatively coarse time resolution. Even where trends in these data have been detected, it has been extremely difficult to attribute them to rising CO₂ levels, because simultaneous changes in many confounding factors — such as rainfall, temperature, land use and fire frequency — have occurred¹¹.

Keenan *et al.* bring a new source of data to bear on this problem. The eddy-covariance technique, developed in the 1980s to quantify the exchange of gases between the atmosphere and land areas, has revolutionized plant-ecosystem science because it continuously monitors the functioning of whole ecosystems on an hourly timescale¹². Using instruments mounted above a vegetation canopy, eddy covariance can be used to measure the carbon uptake and water use of whole ecosystems on a spatial scale of up to one square kilometre. Over the past 20 years, eddy-covariance towers

*This article and the paper under discussion¹ were published online on 10 July 2013.



Figure 1 | Vapour flux. Keenan *et al.*¹ analysed the flux of water vapour and carbon dioxide above forests in the Northern Hemisphere, such as at Willow Creek, California (pictured).

have been established worldwide in a broad range of ecosystems, and high-quality long-term data sets are now becoming available¹².

The authors used these data to analyse long-term changes in ecosystem-scale water-use efficiency. Plants lose water through evaporation whenever they open their stomata to admit CO₂ for photosynthesis, and water-use efficiency is a measurement of the rate at which the plant exchanges water for carbon. This measurement is expected to be a good indicator of the effects of rising CO₂ concentrations on vegetation — because rising CO₂ levels both increase carbon uptake by plants and reduce plant water use, the effects of higher levels of CO₂ on water-use efficiency should be larger and more consistent than the effects on carbon gain or water use alone.

Keenan and colleagues report that the water-use efficiency of forest canopies in the Northern Hemisphere (Fig. 1) over the past two decades shows a remarkable upward trend. The trend was consistent, with no decreases at any

of the 21 forest sites examined. The rates of increase across all sites were large (averaging about 3% per annum) and highly statistically significant.

To prove that this rise was caused by increasing levels of atmospheric CO₂, Keenan *et al.* examined a suite of potential confounding factors. They found that, across all 21 sites, there were no trends in meteorological variables (precipitation, wind speed, temperature and humidity) or structural attributes of canopies (leaf area, leaf-nitrogen content and canopy roughness) that could explain the observed rise in water-use efficiency. The authors conclude that the trend was most consistent with a strong fertilization effect of increased CO₂.

This observation-based finding will probably stimulate considerable debate and further research because, although the reported trend is convincing, the magnitude of the trend is much larger than would be predicted from our existing knowledge of plant responses to

CO₂. Decades of controlled experiments^{13–15} have consistently found that the intercellular CO₂ concentration (C_i) in photosynthesizing tissue is proportional to the atmospheric CO₂ concentration (C_a) — that is, C_i/C_a is constant. The trend identified by Keenan *et al.*, however, implies that intercellular CO₂ has remained constant, and so C_i/C_a has strongly decreased with increasing concentrations of atmospheric CO₂.

To put it another way, controlled experiments^{16,17} have found that the effect of increased atmospheric CO₂ levels on water-use efficiency is roughly proportional to the increase in atmospheric CO₂. By contrast, according to our calculations, the increase in water-use efficiency found by Keenan and co-workers in the eddy-covariance data is approximately six times larger than the corresponding increase in atmospheric CO₂. Consequently, the authors show that current models of the terrestrial carbon cycle do not capture the magnitude of the trend obtained from eddy covariance. This is to be expected, because the models were developed from, and so are consistent with, data from controlled experiments¹⁷.

Keenan and colleagues' study thus provides an intriguing challenge to our understanding of ecosystem functioning — there is a significant increase in water-use efficiency that we cannot currently explain. The implication is either that plants are markedly more responsive to rising CO₂ levels than we previously thought, or that there are other, unknown factors behind the observed trend in the eddy-covariance data. Our view is that it is unlikely that the effect of CO₂ on water-use efficiency is being underestimated by the magnitude that Keenan *et al.* suggest, because the response of water-use efficiency to CO₂ is so predictable in controlled experiments^{16,17}. Nevertheless, the authors have ruled out most other potential drivers for the trend. More research is clearly needed to understand these findings, including long-term studies with complementary observational data streams, such as measurements of changes in plant biomass and water use by whole ecosystems. ■

Belinda Medlyn and Martin De Kauwe
are in the Department of Biological Science,
Macquarie University, North Ryde, New South
Wales 2109, Australia.
e-mails: belinda.medlyn@mq.edu.au;
mdekauwe@gmail.com

1. Keenan, T. F. *et al.* *Nature* **499**, 324–327 (2013).
2. www.esrl.noaa.gov/gmd/ccgg/trends
3. Eamus, D. & Jarvis, P. G. *Adv. Ecol. Res.* **19**, 1–55 (1989).
4. Arora, V. K. *et al.* *J. Clim.* <http://dx.doi.org/10.1175/JCLI-D-12-00494.1> (2013).
5. Lewis, S. L., Malhi, Y. & Phillips, O. L. *Phil. Trans. R. Soc. Lond. B* **359**, 437–462 (2004).
6. Gedney, N. *et al.* *Nature* **439**, 835–838 (2006).
7. Piao, S. L. *et al.* *Proc. Natl Acad. Sci. USA* **104**, 15242–15247 (2007).
8. Wright, S. J. *Glob. Change Biol.* **19**, 337–339 (2013).

9. Clark, D. A., Clark, D. B. & Oberbauer, S. F. *J. Geophys. Res.* **118**, 1–12 (2013).
10. Silva, L. C. R. & Anand, M. *Glob. Ecol. Biogeogr.* **22**, 83–92 (2013).
11. Donohue, R. J., McVicar, T. R. & Roderick, M. L. *Glob. Change Biol.* **15**, 1025–1039 (2009).
12. Baldocchi, D. *Aust. J. Bot.* **56**, 1–26 (2008).
13. Wong, S. C., Cowan, I. R. & Farquhar, G. D. *Plant Physiol.* **78**, 821–825 (1985).
14. Drake, B. G., González-Meler, M. A. & Long, S. P. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **28**, 609–639 (1996).
15. Ainsworth, E. A. & Rogers, A. *Plant Cell Environ.* **30**, 258–270 (2007).
16. Barton, C. V. M. *et al. Glob. Change Biol.* **18**, 585–595 (2012).
17. De Kauwe, M. G. *et al. Glob. Change Biol.* **19**, 1759–1779 (2013).

RNA BIOPHYSICS

A three-state balancing act

How do pathogens survive temperature variations? At a molecular level, one bacterial species seems to regulate gene expression in response to temperature through structural equilibria in corresponding RNA sequences. [SEE LETTER P.355](#)

RONALD MICURA

In bacteria, many messenger RNA molecules carry a regulatory segment called a riboswitch. Specific binding of small ligand molecules, such as adenine, to this segment determine whether the riboswitch mRNA will be translated into a protein^{1–4}. On page 355 of this issue, Reining *et al.*⁵ show that regulation of gene expression through such ‘riboswitching’ is coupled to temperature sensing. The authors investigate the adenine-sensitive riboswitch from the pathogenic bacterium *Vibrio vulnificus*⁶, and demonstrate that efficient RNA regulation at different temperatures — those of the bacterium’s marine habitat and its human host — requires a change in the riboswitch’s structural behaviour from a two-state pattern to a three-state one.*

A typical riboswitch consists of two domains at the 5′ end of the mRNA: a ligand-binding aptamer and an adjoining expression platform, which can have one of two mutually exclusive structures depending on whether the aptamer is in the ligand-bound or ligand-unbound state. The structural change in the expression platform signals that gene expression should be turned on or off.

During gene transcription, a polymerase enzyme synthesizes first the aptamer and then the expression platform⁷. The sequential release of riboswitch domains from the polymerase is especially meaningful for transcription-controlling riboswitches, which act under kinetic control. To direct the folding of the expression platform, the aptamer domain of the growing RNA chain must bind rapidly to its ligand, which requires high ligand concentrations⁸; otherwise, the resulting full-length mRNA becomes trapped in a default fold that cannot respond to the ligand⁹. In

transcription-controlling riboswitches, the two mutually exclusive structures are generally referred to as terminator and antiterminator folds, causing cessation of polymerase activity and continuation of mRNA synthesis, respectively.

Bacterial transcription is tightly coupled to translation. In translation-controlling riboswitches, the molecular mechanism relies on either sequestration or liberation of the Shine–Dalgarno sequence — the mRNA site that binds to cellular organelles known as ribosomes to initiate translation. In contrast to transcription-controlling riboswitches, most translation-controlling riboswitches act under thermodynamic control⁹. Consequently, ligand-dependent control of translation is maintained even for a full-length mRNA. For these riboswitches, therefore, two states (ligand-bound and -unbound) seem sufficient to turn translation on and off.

Reining and colleagues have found that, for robust functioning, their translation-controlling adenine-sensing riboswitch must occur in three structural states. The authors investigated the full-length (more than 100-nucleotide) riboswitch domain at single-nucleotide resolution. They also determined a complete set of thermodynamic and kinetic parameters for folding and ligand-binding of this RNA, under conditions that involved varying the concentrations of RNA, magnesium ions (a factor mediating structure formation) and adenine, as well as, importantly, temperature.

The researchers find that a ligand-free (apo) form of the riboswitch exists in a pre-equilibrium of two structurally distinct aptamer folds (Fig. 1). One of the structures (apoA) can bind to the ligand, exhibiting a structure that resembles a third structural state — the adenine-bound holo form. The other structure (apoB) adopts a different fold and cannot interact with the ligand. But why does



50 Years Ago

A technique has been used with limited success to obtain simultaneous measurement of cosmic radiation at two different altitudes. The method consists of suspending two packets of nuclear emulsion plates from the one balloon, while maintaining a constant vertical separation of 10,000–27,000 ft. between the packets ... in this way the two packets do not separate in latitude and longitude—a factor which enters if two separate independent balloons are flown.

This has been achieved by carrying aloft a cone of nylon string which was allowed to unwind at a predetermined height, leaving one packet of plates suspended at the balloon while the other fell at the end of the string ... As the alarm rings, the alarm winder releases the key ... holding the lower packet and target, which then fall away freely, unwinding the string in the process. Twenty thousand feet of string unwinds in approximately 10 min. **From *Nature* 20 July 1963**

100 Years Ago

The Potato: A Compilation of Information from Every Available Source. By E. H. Grubb and W. S. Guilford.

There are men who, having attained to wealth and fame by the agency of some humble instrument, basely repudiate and kick over the ladder by which they have risen. Not so the authors of the first book on our list. The potato has “made” them, and in return they proceed to “make” the potato ... the authors are so evidently enthusiastic, and discourse so eloquently on the merits of their subject, that we are carried along with them, and forget that, after all, they are only talking about potatoes, and not about alpine plants or roses.

From *Nature* 17 July 1913

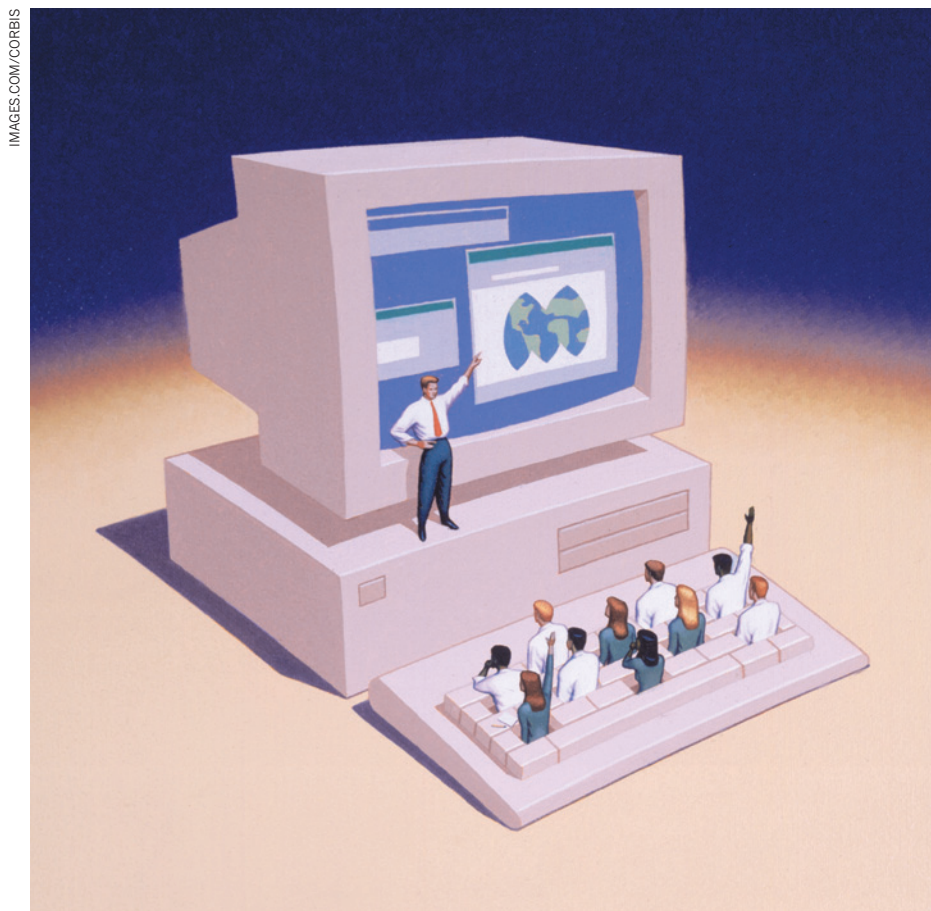
*This article and the paper under discussion⁵ were published online on 10 July 2013.

CAREERS

TURNING POINT Geneticist finds recognition as an innovator and toolmaker **p.371**

NATUREJOBS FACEBOOK Science-job resources, discussions and more go.nature.com/4lxra

NATUREJOBS For the latest career listings and advice www.naturejobs.com



ONLINE LEARNING

How to make a MOOC

With forethought and support, science instructors can design effective massive open online courses.

BY SARAH KELLOGG

When Marnie Blewitt was designing her first online epigenetics course, she knew that she would need help. A geneticist at the University of Melbourne in Australia, Blewitt was new to the business of MOOCs — massive open online courses that

can reach thousands or even hundreds of thousands of students around the globe.

Blewitt enlisted Kylie Greig, a biology postdoc with an aptitude for technology, to help her to construct quizzes, create instructional slides and images and manage online student forums during the six-week course, which launched on 1 July on Coursera in Mountain View, California

— the largest MOOC platform worldwide. Blewitt and Greig will also work together to analyse test data and determine whether they need to add new content to clarify certain principles.

The help was crucial. “It is a very different feeling from a regular university course,” says Blewitt. “The sheer number of students just really adds to the pressure, and also the knowledge that people will post immediately if they don’t like it.”

In the past couple of years, academic institutions, policy centres and research institutes around the world have become partners with MOOC providers including Coursera, edX in Cambridge, Massachusetts, and Udacity in Mountain View (see Comment, page 275). Coursera has so far registered 9.5 million enrolments; edX counts some 1 million. Institutions or instructors looking to go it alone can post lecture videos on YouTube.

MOOCs, which are usually free of charge, have course assignments and exams like their on-campus cousins, but they also feature discussion boards on which tens of thousands of students can gather to share ideas and discuss lectures. Course subjects are as disparate as the students, ranging from ‘Artificial Intelligence for Robotics’ to the ‘History of Rock: Part Two’ (which covers music, not minerals).

Instructors may feel daunted by the challenge of making a popular course shorter, livelier and easily understood on a computer screen. However, experienced MOOC designers say that the task is well worth the effort, because the courses allow for experimentation on a grand scale. Although it does not generally bring in extra pay, MOOC work gives teachers a chance to develop fresh methods for teaching science, bringing it to new, uninitiated audiences and encouraging cutting-edge discussions.

A DIFFERENT WORLD

MOOC creation requires rigorous attention to detail, with long hours of preparation and instruction. Successful courses are more than a lecture uploaded to the Web, say experts, and they serve a broader and more varied audience than other types of distance learning.

“Instructors designing MOOCs have to think through the whole new medium,” says Anant Agarwal, president of edX and a computer scientist at the Massachusetts Institute



DIGITAL LEARNING

A *Nature* and *Scientific American* special. nature.com/digitalllearning

► of Technology (MIT) in Cambridge. “How can materials be given to students in bite-sized chunks? You cannot stand at a board and talk for an hour, because students will lose interest, and you have to use discussion boards to guide student conversations about the material.”

MOOCs also differ from on-campus teaching in that students do not necessarily learn the same thing at the same time. Udacity makes its courses available all year round, so that students can work at their own pace, and MOOCs on YouTube or other online-learning sites may be available in perpetuity, although instructors or universities often take down outdated material. By contrast, edX and Coursera offer courses at specific scheduled times, to ensure that student discussion boards and forums are monitored, and that all students are at the same point in their studies.

One of the benefits of MOOCs is the ability to make frequent updates to reflect changes in technology or the thinking on a subject. Eric Lander, a biologist at MIT, teaches an introductory biology class that he has adapted into a MOOC for edX. He taught it for the first time in March, and will run it again in September. But before that, he plans to revise it, adding new video lectures to reflect lessons learned during the first course. “We asked ourselves, ‘If we have access to Web-based media, how would we do the course differently for online students?’” says Lander. For his team, the answer was to inject technologies that they had not used before, namely animations and three-dimensional models.

Another prized advantage of MOOCs is the discussion boards. These need to be guided and monitored, but they can host intense debates and thought-provoking dialogues about the subject matter, the course and even the instructor. And active boards encourage students to rate peer comments, often creating a chatting frenzy. They also allow for peer-to-peer learning, with students, for example, coaching each other on homework assignments or helping to review each other’s quiz results.



Graduate student Brian Lovett helps teacher Raymond St. Leger set up for a palaeogenomics lecture.

When it comes to designing a MOOC, instructors recommend that first-timers recruit a support team — perhaps including curriculum-development specialists, computer programmers, videographers or tech-savvy teaching assistants — to help them to clear pedagogical and technical hurdles (see ‘Get started in MOOCs’). “It takes a village,” says Lander. “I’ve taught my course for 23 years, and I know the material, but I needed help to convert it to a new platform, so it was good to have the outside expertise,” he says. On campus, his course enrolls as many as 850 students at a time; online, it drew some 35,000 students.

STANDING OUT IN THE CROWD

Most MOOC companies offer comprehensive services to help instructors to design their classes, assess student learning and create assignments. Some instructors think that these systems make the courses more generic by, for example, removing the distinct voice of the teacher. But many believe that the benefits outweigh the downside: sacrificing some personal lecturing style allows them to simplify the MOOC-creation process and gain the input of

technology experts. Practised instructors say that teachers should think about what they want the students to learn and experience, rather than letting the platforms drive the content.

Instructors should also rethink their lecture styles, even down to their body language. Enthusiastic teaching with broad gestures and loud voices may add drama in a lecture hall, but it is a distraction on a small screen. Veteran instructors recommend that novices record lectures in front of students in a lecture hall or in smaller groups, and that they take time to acclimatize to the technology — which can include high-tech cameras and green screens for inserting different background. “I had to learn to be more still because the camera picks up movement,” says Raymond St. Leger, an entomologist at the University of Maryland in College Park, who taught a MOOC on Coursera in April. He worried about giving a leaden performance after lots of rehearsing. “I had to get used to teaching to a blinking light on the camera and learn how to be more spontaneous.”

“Overall, I think there’s an optimal mix of entertainment and instruction that keeps students interested while still being useful for learning the material,” says Michael DeWeese, a physicist at the University of California, Berkeley, who has taught an introductory physics class as an online course. It is now posted on YouTube.

KNOW YOUR AUDIENCE

It is crucial to consider who the students will be. Some may have degrees in the discipline, but others will have little or no experience. Veteran instructors say that it is essential to provide background information that lays out basic principles and knowledge early in the course.

“Online students are different from students on campus,” says Michael Evans, a chemist at the University of Illinois at Urbana-Champaign, who has worked with his PhD supervisor, Jeffrey Moore, to design and teach an organic-chemistry course on Coursera. “Online students are older and they’re looking for real-world

COURSE DESIGN

Get started in MOOCs

- Reach out to experts in curriculum development and technology for guidance.
- Develop a list of learning goals that can be accomplished in brief video segments or through methods such as readings, discussion boards and guest lectures.
- Don’t dumb it down, but do prepare background materials for non-scientists.
- Film lab experiments, create interactive tools and slides, and develop problem sets rendered with three-dimensional graphics.
- Try out videos, online quizzes and animated models on a small cohort first.
- Students may access material at different times and out of order, so avoid mentions of lecture dates, times and content sequence.
- Practice the lecture before taping.
- Hire an assistant to help in monitoring forums and distributing communications.
- Make course expectations regarding accessing homework and quizzes clear early on, while also mapping out deliverables.
- Encourage student-to-student interaction and mentoring in online forums.
- Collect data on student learning to improve the next iteration of the course. **S.K.**

applications for what they're learning."

Evans says that instructors need to be imaginative with their lectures because they do not have forces such as peer pressure keeping students in seats, as they would in the classroom. "People have different ways of teaching things, but with video, you only have one chance to say something," says Evans, noting that a failure to keep the online class engaged can result in hundreds, if not thousands, of students logging off. "There are no tangents or wandering off. You get a snapshot in time, and it doesn't allow for even good digressions." He suggests avoiding anecdotes and verbal detours that make sense in the classroom but might not work online.

COURSE UNCERTAINTIES

Some sceptics remain doubtful that MOOCs can foster a satisfactory learning experience, no matter what technologies are used. They worry that students do not acquire the same breadth of knowledge as in a classroom-based course, and that potentially important bits of content can be left out. Efforts such as discussion boards do not really create a sense of community that replicates a classroom, says Carl Wieman, director of the Carl Wieman Science Education Initiative at the University of British

Columbia in Vancouver, Canada. "I wouldn't deny that's possible," he says. "I would deny that anybody has demonstrated they have technology that can do it."

Concerns such as these can be mitigated by community-building technologies that allow students to build connections with each other and their teachers.

Instructors can offer one-on-one discussions using Web-based communications programmes; they can also get involved in interactions on discussion forums.

For early-career instructors, the broad appeal of MOOCs can offer something else: name recognition. "Younger faculty do see it as an opportunity for fame, if not fortune," says Evans. "You can really get your name out there far faster." That means that the stakes are high. "We taught more students in this one course," says St. Leger, "than we probably will teach in person for the rest of our careers." ■

Sarah Kellogg is a freelance writer in Washington DC.

TURNING POINT

Thijn Brummelkamp

Thijn Brummelkamp will receive a Gold Medal and €10,000 (US\$13,000) in September from the European Molecular Biology Organization (EMBO), based in Heidelberg, Germany, for co-developing techniques that allow genes in human cells to be inactivated. A geneticist at the Netherlands Cancer Institute (NKI) in Amsterdam, Brummelkamp also co-founded Haplogen, a company in Vienna that aims to identify targets for drugs that treat infectious diseases. He details how he came to be known as an innovator and toolmaker.



How did you become interested in genetics?

When I was young, I liked taking apart radios and televisions. DNA comes close to that: you can break it down and see the consequences. When I started my PhD at the NKI, my project was to identify genes that play a part in cancer. I became frustrated by the limited options for studying the genetics of human cells as opposed to those of model organisms. You cannot do genetic crosses in human cells.

What did you do to tackle this problem?

A colleague and I made a DNA construct to produced hairpin-shaped RNA molecules that silenced genes indefinitely. In the two or three months after we published the research in 2002, about 1,500 people asked for our reagents. You expect a few reactions, but not 1,500. More than 10 years later, people still use hairpin RNAs to examine gene function in human cells. You can make libraries of these hairpin RNA molecules, silence many genes and see which gene is linked to a phenotype of interest.

What did you do after your PhD?

I became a Whitehead Fellow at the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts. The fellowship allows you to work for several years as a group leader without the stress of a tenure clock. And it gives you the freedom to find and start projects.

In 2005, you were named one of the top innovators under the age of 35 by MIT Technology Review. How did you feel?

I'd heard of the list, but it was a surprise. It's wonderful to be recognized so early in your career. That can help to convince people to come to your lab, and that is important in the beginning.

What challenges did you face when you moved to the NKI in 2010?

Moving a lab definitely involves some logistics. It takes years to set yourself up — you need to train people, and write grants and

safety protocols. If you switch continents, suddenly it all needs to be done again. Not all my lab personnel joined me, so I lost people with experience. Also, transferring grants is not always easy. Almost all animal and safety protocols are different: for example, virus safety-level categorizations can differ from country to country. This slowed things down a bit. I had to wait before experiments could be done, and some experiments that could not have been performed here were done by collaborators in the United States. But now things are fine.

How did co-founding Haplogen affect your career?

It's one of the most exciting things I've ever been involved in. You meet entirely different people — people who work in companies and who help with financing. I'm learning more about the process of drug development, and that's an entirely different world.

What was your reaction when you won the EMBO Gold Medal?

I had not expected I would be considered, let alone that I'd receive it. It's wonderful because many people have noticed it, and I've already had some invitations to seminars and meetings because of it. It's a bit frightening because the expectations might be difficult to live up to.

Do you have any advice for graduate students and postdocs?

What has helped me is being involved in making tools. We could look at known biological problems with new techniques. If you make the tool yourself, you have first access and the most experience with it. That has kept us in business. ■

INTERVIEW BY ROBERTA KWOK



"It is a very different feeling from a regular university course."

Marnie Blewitt

applications for what they're learning."

Evans says that instructors need to be imaginative with their lectures because they do not have forces such as peer pressure keeping students in seats, as they would in the classroom. "People have different ways of teaching things, but with video, you only have one chance to say something," says Evans, noting that a failure to keep the online class engaged can result in hundreds, if not thousands, of students logging off. "There are no tangents or wandering off. You get a snapshot in time, and it doesn't allow for even good digressions." He suggests avoiding anecdotes and verbal detours that make sense in the classroom but might not work online.

COURSE UNCERTAINTIES

Some sceptics remain doubtful that MOOCs can foster a satisfactory learning experience, no matter what technologies are used. They worry that students do not acquire the same breadth of knowledge as in a classroom-based course, and that potentially important bits of content can be left out. Efforts such as discussion boards do not really create a sense of community that replicates a classroom, says Carl Wieman, director of the Carl Wieman Science Education Initiative at the University of British

Columbia in Vancouver, Canada. "I wouldn't deny that's possible," he says. "I would deny that anybody has demonstrated they have technology that can do it."

Concerns such as these can be mitigated by community-building technologies that allow students to build connections with each other and their teachers.

Instructors can offer one-on-one discussions using Web-based communications programmes; they can also get involved in interactions on discussion forums.

For early-career instructors, the broad appeal of MOOCs can offer something else: name recognition. "Younger faculty do see it as an opportunity for fame, if not fortune," says Evans. "You can really get your name out there far faster." That means that the stakes are high. "We taught more students in this one course," says St. Leger, "than we probably will teach in person for the rest of our careers." ■

Sarah Kellogg is a freelance writer in Washington DC.

TURNING POINT

Thijn Brummelkamp

Thijn Brummelkamp will receive a Gold Medal and €10,000 (US\$13,000) in September from the European Molecular Biology Organization (EMBO), based in Heidelberg, Germany, for co-developing techniques that allow genes in human cells to be inactivated. A geneticist at the Netherlands Cancer Institute (NKI) in Amsterdam, Brummelkamp also co-founded Haplogen, a company in Vienna that aims to identify targets for drugs that treat infectious diseases. He details how he came to be known as an innovator and toolmaker.



How did you become interested in genetics?

When I was young, I liked taking apart radios and televisions. DNA comes close to that: you can break it down and see the consequences. When I started my PhD at the NKI, my project was to identify genes that play a part in cancer. I became frustrated by the limited options for studying the genetics of human cells as opposed to those of model organisms. You cannot do genetic crosses in human cells.

What did you do to tackle this problem?

A colleague and I made a DNA construct to produced hairpin-shaped RNA molecules that silenced genes indefinitely. In the two or three months after we published the research in 2002, about 1,500 people asked for our reagents. You expect a few reactions, but not 1,500. More than 10 years later, people still use hairpin RNAs to examine gene function in human cells. You can make libraries of these hairpin RNA molecules, silence many genes and see which gene is linked to a phenotype of interest.

What did you do after your PhD?

I became a Whitehead Fellow at the Whitehead Institute for Biomedical Research in Cambridge, Massachusetts. The fellowship allows you to work for several years as a group leader without the stress of a tenure clock. And it gives you the freedom to find and start projects.

In 2005, you were named one of the top innovators under the age of 35 by MIT Technology Review. How did you feel?

I'd heard of the list, but it was a surprise. It's wonderful to be recognized so early in your career. That can help to convince people to come to your lab, and that is important in the beginning.

What challenges did you face when you moved to the NKI in 2010?

Moving a lab definitely involves some logistics. It takes years to set yourself up — you need to train people, and write grants and

safety protocols. If you switch continents, suddenly it all needs to be done again. Not all my lab personnel joined me, so I lost people with experience. Also, transferring grants is not always easy. Almost all animal and safety protocols are different: for example, virus safety-level categorizations can differ from country to country. This slowed things down a bit. I had to wait before experiments could be done, and some experiments that could not have been performed here were done by collaborators in the United States. But now things are fine.

How did co-founding Haplogen affect your career?

It's one of the most exciting things I've ever been involved in. You meet entirely different people — people who work in companies and who help with financing. I'm learning more about the process of drug development, and that's an entirely different world.

What was your reaction when you won the EMBO Gold Medal?

I had not expected I would be considered, let alone that I'd receive it. It's wonderful because many people have noticed it, and I've already had some invitations to seminars and meetings because of it. It's a bit frightening because the expectations might be difficult to live up to.

Do you have any advice for graduate students and postdocs?

What has helped me is being involved in making tools. We could look at known biological problems with new techniques. If you make the tool yourself, you have first access and the most experience with it. That has kept us in business. ■

INTERVIEW BY ROBERTA KWOK



"It is a very different feeling from a regular university course."

Marnie Blewitt